AFOSR·TR· 89·1735 ②

University of Idaho

The Formation and Use of Knowledge

Structures in Problem Solving Domains

Sallie E. Gordon
Department of Psychology

Richard T. Gill
College of Engineering

90   01 04 041

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 89-1735

6a. NAME OF PERFORMING ORGANIZATION University of Idaho	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research/NL
6c. ADDRESS (City, State, and ZIP Code) Moscow, ID 83843		7b. ADDRESS (City, State, and ZIP Code) Building 410 Bolling AFB, DC 20332-6448

8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR	8b. OFFICE SYMBOL (If applicable) NL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-88-0063

8c. ADDRESS (City, State, and ZIP Code) Building 410 Bolling AFB, DC 20332-6448	10. SOURCE OF FUNDING NUMBERS			
	PROGRAM ELEMENT NO. 61102F	PROJECT NO. 2313	TASK NO A4	WORK UNIT ACCESSION NO

11. TITLE (Include Security Classification)

The Formation and Use of Knowledge Structures in Problem Solving Domains

12. PERSONAL AUTHOR(S)
Sallie E. Gordon & Richard T. Gill

13a. TYPE OF REPORT Final Technical Report	13b. TIME COVERED FROM 1/1/88 TO 9/30/89	14. DATE OF REPORT (Year, Month, Day) October 4, 1989	15. PAGE COUNT 63

16. SUPPLEMENTARY NOTATION

17 COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	memory, problem solving, declarative knowledge structures, individual differences
05	10		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The goal of this research was to develop and test a method for eliciting knowledge structures used in problem solving. The work was carried out simultaneously in two domains, engineering mechanics and video recording. Two studies resulted in the adaptation of a question probe method for eliciting relevant knowledge structures prior to problem solving. Two additional studies showed that administration of the question probes did not significantly impact subsequent problem solving. Answers from the question probes were therefore translated into a conceptual structure for each subject. A simple associative search model operating upon these structures was able to predict 87% and 93% of individual problem solving activities in the engineering and video recording domains, respectively. The predictive capability of the graphs indicate the central role of knowledge structures in problem solving processes, at least under circumstances such as those tested. A fifth study evaluated the validity of the conceptual graphs by comparing them with free recall protocols. Content and clustering patterns in free recall data were highly related to the content and organization of the question probe-based graphs, lending support for the assumption that the graphs reflect underlying declarative knowledge structures.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Alfred R. Fregly	22b. TELEPHONE (include Area Code) (202) 767-5021	22c. OFFICE SYMBOL NL

DD FORM 1473, JUN 86
Previous editions are obsolete
SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED   16 OCT 1989

Summary

Five studies were conducted to develop and evaluate a method for eliciting knowledge structures used in problem solving. The work was carried out simultaneously in two problem domains, engineering mechanics and video recording. Two studies, one in each domain, demonstrated that a question probe method could be successfully adapted from work in prose comprehension to elicit individual knowledge structures prior to problem solving. Two additional studies tested the intrusiveness of question probes: Subjects watched an instructional videotape and then either were or were not administered a large number of question probes before solving problems. Data analysis showed that administration of question probes did not significantly affect subjects' problem solving performance in either domain.

Because of the lack of intrusiveness, it was possible to study the relationship between knowledge structures and problem solving on an individual basis. Question probe answers were used to construct a conceptual structure or graph for each subject. Completeness and proportional accuracy of these graphs were highly predictive of problem solving scores (R = .88 and .82 for engineering mechanics and video recording, respectively).

In addition, a problem solving model assuming the operation of a simple spreading activation search mechanism was used to predict problem solving behavior. More specifically, the content of each graph was used to derive predictions for whether that subject would successfully perform or fail to perform each of the subtasks or procedures necessary to solve the problems. The model predictions matched 87% of subject actions in the engineering mechanics domain and 93% of subject actions in the video recording domain. Because of certain mitigating factors, the prediction rate for engineering mechanics was considered to be a conservative estimate.

The fifth study evaluated the accuracy or validity of conceptual graphs developed using question probes. Subjects were given both free recall and question probe tasks after watching an instructional tape. Clustering patterns in the free recall data were highly related to the organization of the conceptual graphs. This lends support to the validity of the question probe method and resulting graphs.

Taken together, these studies strongly suggest a central role of knowledge structures in problem solving. The content of the structures accounted for a great majority of the variance in performance scores and also for almost all of the specific actions taken. The fact that conceptual structures developed with question probes can be highly predictive of individual performance makes the approach very promising for future study of knowledge structure and process in problem solving.

Table of Contents

INTRODUCTION

Objectives

The objectives of the research project were to develop and evaluate a method to elicit declarative knowledge structures and study their use in problem solving. Since a promising method from the area of prose comprehension had been identified, the specific goals of the project were to (1) adapt the question probe method to the area of problem solving, (2) test the feasibility of using the method in two different types of problem domain, (3) evaluate the effects of administering question probes to subjects on subsequent problem solving, (4) use the question probe method to evaluate the use of knowledge structures in problem solving, and (5) collect preliminary data on the validity of the knowledge structures constructed from question probes.

Background

Traditionally, researchers in the area of reasoning and problem solving have primarily relied on the method of "think aloud" verbal protocol analysis (Ericsson & Simon, 1984; Payne, 1980). The assumption is that the protocols can provide an indication of the strategies being tried, the cognitive "procedures" being used, and so on. Based on these verbal protocols, models are built of the procedures which subjects are presumably using in problem solving. There are inherently several problems with this approach including: (1) The models are based on very incomplete data, that is, only information brought to consciousness during problem solving; and (2) The problem solving protocols are used to develop the models and also as validation for the models. However, the most serious problems is: (3) The models are only very general heuristics (e.g., difference reduction methods, means-end analysis, etc.) and cannot predict <u>specific</u> problem solving activity and success on significantly different problem.

Recently, several researchers have suggested a need for models detailing the structure and use of knowledge representations in problem solving, in addition to descriptions of the "solution path" followed (e.g., Kintsch & Greeno, 1985). In order to clarify this distinction, consider the following analogy. If a person were traversing a forest, a camera on that person's shoulder would certainly lend a description of which paths were chosen along the way. This could then be drawn as a path with directions, turns, etc. However, this might not tell us anything about how the person would traverse another section of the forest. On the other hand, if one had a map of the entire forest, along with rules for how the person interacted with the various characteristics of that forest, then one could better predict how a person would traverse the forest from any starting point.

This analogy suggests the need for two important characteristics of a predictive model. The first is an understanding of the "forest," or the person's conceptual

3

structures which they are operating upon during the course of problem solving. The second important aspect of the model concerns procedural assumptions or how the conceptual structures are used. If we had an adequate model of what knowledge is used and how it is used, we could theoretically _predict_ problem solving on new problems.

Some researchers have begun investigating the role of knowledge representations in specific domains of reasoning and problem solving. For example, diSessa (1983) and Larkin (1983) developed models of knowledge structures used in the domain of physics, Forbus (1981) suggested a model of knowledge representation for physical systems (Qualitative Process Theory), and Sembugamoorthy and Chandrasekaran (1986) developed a detailed model of a functional representation of a household buzzer. Pennington (1987) studied the type of knowledge structures used by expert computer programmers, and Kintsch and Greeno (1985), Riley, Greeno, and Heller (1981), and Reed (1987) developed models of knowledge representation and use in solving standard verbal mathematics problems. The main drawback of these models is that they are domain specific, and it is not easy to see how they would be extended to other problem solving domains.

Other researchers have taken a more generic approach. For example, Kolodner and Simpson (1986) have suggested that knowledge consists of generalized episodes which are composites of problem solving experiences from the past, where the specific experiences are indexed from the generalized episodes. Rumelhart and Norman (1981) postulate the existence of declarative schemas and procedural schemas where the declarative schemas are built as a _function_ of the procedural schemas. Finally, Stevens and his colleagues (Stevens, Collins, and Goldin, 1979; Williams, Hollan, & Stevens, 1983) have argued that subjects have multiple mental models which are brought to bear on a problem where the models are successively generated by consideration of relevant data, inferences, and inconsistencies in reasoning. However, all of these approaches are incomplete with regard to specification of knowledge representation structure, operations upon those structures, and how new problem representations interact with prior knowledge.

In summary, if we accept the need for representational models specifying the characteristics of knowledge structures and the processes by which they are used, we can then identify several problems with the existing research methodologies. Models developed on the basis of verbal problem solving protocols are incomplete because of the limited source of data and circular evolution and testing of the models. Models based on more complete specifications of knowledge structures are often domain-specific and difficult to generalize to other problem domains.

Our supposition in undertaking this work was that research methods and some aspects of representational models currently being developed in the area of prose comprehension could provide a new and useful approach in the study of reasoning and problem solving. More specifically, models of internal knowledge representation are important to understanding processes in

4

problem solving, and techniques from prose comprehension are available for building these models.

The work reported here is an application of one such technique developed by Graesser and colleagues (Greasser and Clark, 1985; Graesser and Goodman, 1985). Before describing the knowledge structure representational method, we will briefly describe the assumptions underlying the prose comprehension model.

In the Graesser model, it is hypothesized that knowledge structures are constructed at the time a passage is read or heard. A knowledge structure consists of statement nodes which are basic idea units similar to the "propositional schema" used by van Dijk and Kintsch (1983). An example of such a node might be "drank a glass of wine." Each statement node is assigned a node category such as states, events, goals, intentional actions, or style. The statement nodes are interrelated by directed arcs or links. The arcs are categorized as being one of nine types: Consequence, Implies, Reason, Outcome, Initiate, Manner, Property, Set Membership, or Referential Pointer.

The passage conceptual structure is successively built during the course of comprehension. The statements in the structure consist of idea units explicitly contained in the passage, as well as _inferences_ which have been supplied by generic knowledge structures (GKS). GKSs are rich structures of knowledge which are activated when a concept or pattern in the passage matches or maps onto information in long-term memory.

Knowledge representations are divided into the two commonly found classifications of declarative knowledge structures and procedural knowledge structures. Delcarative knowledge structures have the form of the previously described conceptual graph structures. They include specific passage conceptual structures as well as the GKSs. Procedural knowledge consists of active symbolic procedures which _operate upon_ the declarative knowledge structures (these procedures include search operations, matching operations, and structure modification). Notice that this use of "procedural" is different from the concept of procedural in problem solving domains (e.g., Rumelhart and Norman, 1981) where procedural knowledge refers to specific actions to be performed, or "knowledge about procedures".

The Question Probe Method

In order to develop a model of knowledge representation in prose comprehension, Graesser devised a unique technique to extensively probe subject knowledge structures. This "question answering" technique, consisted of using a systematic questioning technique to elicit conceptual structures generated during passage comprehension. Graesser hypothesized that the conceptual structure formed during prose comprehension consisted of statement nodes and arcs corresponding to specific information given in the material, as well as statement nodes and arcs which were _inferences_ provided by GKSs accessed during comprehension. To determine the inferences made during comprehension, Graesser used a question answering technique after subjects had read a

5

particular passage. That is, for each explicit event, state, and intentional action in the passage, subjects were probed with three types of questions: Why? How? and What-happened-next (WHN)?. By using all answers given by two or more subjects, an aggregate conceptual structure was built consisting of the information explicitly provided by the passage as well as the inferences commonly generated by subjects.

Method Adaptation

The long-term goals of our research program are to (1) adapt the question answering method to problem solving domains; (2) develop a theoretical model of knowledge representation in problem solving; and (3) test key assumptions of the theory using both question probes and traditional experimental methods (e.g., recognition reaction time). Thus, the program depends heavily on the successful adaptation of the question answering technique to problem solving domains. The primary goal of this particular project was therefore to adapt the question probe method to problem solving. In addition, given successful adaptation of the method, a secondary goal was to use question probes to begin evaluating the relationship between knowledge structures and problem solving.

The first consideration in adaptation of the technique was an assessment of the differences between prose comprehension and problem solving. One way of viewing this difference is to start with the assumption that prose comprehension is the use of prior general knowledge structures or schemas to aid in forming a "static" episodic representation of the prose passage (see left side of Table 1). Problem solving includes a similar process when instructions in the domain are received (which is the case only in some types of problems), and also when the problem itself is given to or formulated by the problem solver. It can be seen in the table that although there are processes in problem solving that are functionally similar to prose comprehension, there is a critical additional process. This is the process of _operating upon_ the problem representation and other relevant represenations (i.e., the instructional knowledge structure and/or GKSs) to determine the problem solution.

We determined that the study of instruction representation and problem representation could probably be handled in a manner similar to prose comprehension. That is, after instructional information is presented to subjects, it could be followed by question probes appropriate for the particular domain. It was expected that the resulting data would be similar in form to that found in prose comprehension.

Likewise, problem statements could be given to subjects, followed immediately by a set of question probes. These probes could take at least two forms depending on the problem statement. First, the standard set of probes (Why, How, What, etc.) could be asked for each element or clause in the problem statement. However, for some problem statements this would result in peculiar and irrelevant information. For example, in the problem "Arrange ten Christmas trees in five straight rows of four trees

6

Table 1. Comparison of Processes in Prose Comprehension
and Problem Solving.

Prose Comprehension	Problem Solving
Form general schemas	Form general schemas
Receive prose passage	Receive passage "I" (instruction in problem domain)
Combine passage information with inferences from schemas	Combine passage "I" with inferences from general schemas
	Receive passage "P" (problem statement)
	Combine passage "P" with inferences from general schemas and also instruction knowledge structure ("I")
	Operate upon passage "P" knowledge structure and other relevant structures to solve problem
	(Above process will result in new "static" structures)

each," the experimenter could ask: "Why are there christmas trees? How are there christmas trees? and What are christmas trees?." It can be seen that this set of questions would sometimes result in odd or irrelevant answers. Thus, one of the first tasks was to determine what question probes would be appropriate in problem solving domains. A related question was whether a different set would have to be identified for each different domain or whether there was some degree of overlap.

Besides the issue of identifying appropriate question probes, there were several other considerations in adapting the method to problem solving. For example, Graesser developed an aggregate conceptual graph describing the answers of most subjects combined. This graph automatically included all statements presented in the story. This approach did not seem particularly applicable to the study of knowledge structures used in problem solving for two reasons. First, we cannot assume all subjects completely and accurately represent information from

7

instructional materials. Second, it would be desirable to represent individual conceptual structures which could then be related to individual problem solving performance. For these reasons, only the essence of the question answering methodology itself was retained and used as a starting point for this research.

Research Overview

Five studies were conducted to accomplish the research goals. Experiments 1 and 3 were conducted to develop instructional materials and identify appropriate question probes and problems in the domains of engineering mechanics and video recording. Experiments 2 and 4 were conducted to evaluate the intrusiveness of question probes in the two domains, and use question probe data to predict subject problem solving performance. Experiment 5 validated the use of question probes for eliciting knowledge structures by comparing the organization of the structures with free recall clustering.

In all experiments, subjects viewed an instructional videotape describing the basic concepts and procedures of the domain. In all studies except Experiment 5, subjects were then given some combination of question probes and problems to solve.

For the engineering mechanics (EM) domain, naive subjects were taught three-dimensional vector analysis. For problems, they were given graphical representations of vectors in two and three-dimensions, and asked to mathematically model the vectors.

For the video recording (VCR) domain, subjects were allowed to view a state-of-the-art Hi-Fi, flying head, video recorder while watching an instructional videotape on how to use the equipment. This piece of equipment is new on the market and has numerous highly specialized functions and characteristics not commonly found on video recorders. Subjects were then asked to perform several tasks requiring the use of multiple Hi-Fi VCR functions.

The domains of engineering mechanics and video recording were chosen for a variety of reasons. It was felt that the technique should be tried in problem domains that varied in terms of domain-specific versus general knowledge structure use. It seemed that the EM problems would require primarily domain-specific knowledge, while the VCR problems would require a mix of both new domain-specific knowledge and previously existing generic knowledge about VCRs. Furthermore, these two problem domains are sufficiently different to enhance the generalizability of any findings.

EXPERIMENT 1

The primary goal of Experiment 1 was to develop and finalize materials in the domain of engineering mechanics. The specific goals were (1) to elicit feedback from subjects regarding the instructional tape in order to rectify any problems, (2) to generate and evaluate a wide set of question probes relevant to the instructional material and to the problem

statements, (3) to obtain data on a large and varied set of problems in order to choose a range for use in Experiment 2, and (4) obtain subject feedback on the clarity, difficulty, etc. of the problems in order to make any necessary changes. Although this was essentially a pilot study, the process of developing question probes was central to the project and the study will therefore be described in detail.

Method

Design. Subjects were screened via a mathematics background questionnaire, viewed an instructional videotape on mathematically writing vectors, and then answered question probes and solved problems. Subjects were randomly assigned to one of four conditions:

(1) Instructional Question Probes - subjects received question probes for the instructional material then solved three problems.

(2) Problem Question Probes - for each of the three problems, subjects received a problem statement, were given question probes for the problem statement, and then solved the problem.

(3) Instructional and Problem Question Probes - subjects received the question probes on the instructional material then received the same sequence of tasks as subjects in group 2.

(4) Instructional Question Probes After - subjects worked all three problems before answering question probes for all the instructional material.

Subjects. Subjects were 24 male and female University of Idaho Introductory Psychology students participating for course credit and a $2 cash bonus for each problem solved correctly.

Materials. A subject background questionnaire was developed to screen subjects for the mathematical background necessary to understand the material (college algebra) but no knowledge of writing vectors per se. The questionnaire requested information on age, major, year in college, and previous math courses.

To develop the instructional tape, a preliminary 40-minute instructional videotape was filmed. The content consisted of black and white written statements (essentially an extended outline), equations, and drawings accompanied by verbal explanations equivalent to a classroom lecture on the topic. The material in the videotape was then transformed into a conceptual graph by two members of the research team.[1] The graph revealed several missing concept relationships and one inconsistency in the material. The instructional material was revised and a new videotape was produced, lasting approximately 45 minutes.

Two sets of question probes were created, a set to map

---

[1] The conceptual graphs were composed of node/link/node segments. Following Kintsch (1988), a node could be either a simple concept or a proposition. While not using Sowa's (1984) formalism for writing graphs, our use of the terms "conceptual structure" and "conceptual graph" is not inconsistent with his work.

subjects' representations of the instructional material (hereafter referred to as instructional question probes) and a set to map subjects' representations of the problems (hereafter referred to as problem question probes).

The first step in developing the instructional question probes was to transform the instructional tape into a conceptual graph. Appendix I lists the node-link-nodes in the conceptual graph (listed in the left column of the appendix). The next step was to identify all nodes in the conceptual graph and then generate sets of questions (what, why, etc.) for each node. Unfortunately, it was immediately apparent that following such an approach would result in an enormous number (well over 500) of question probes. We felt that this would undoubtedly result in fatigue and impatience in our subjects. To resolve this problem, we decided to restrict the type of question probes to be used.

We first determined that for each node in the instructional materials, subject's conceptual structures would potentially contain two types of links; links reflecting material explicitly given or strongly implied in the tape, and links that had been formed by subjects as a result of inferences. We decided to constrain our question probes to only those related to links given or implied in the instructional tape. For example, if a node had adjoining links of the type PROPERTY, a corollary question would be given such as "What are the properties of X?" However, if a node concept had no properties provided in the instructional tape, no property type of probe would be given. This resulted in a set of 72 question probes. These question probes are listed on the right side of Appendix I and include questions listed in parentheses. The majority of the question probes were one of the following types; what is __, what are the properties of __, how do you __ (or what is the procedure for __), what is the result or consequence of__, what is the equation for __, and why do you __?

To elicit representations for each problem statement, one general question was used, "describe the problem as completely as you can." If the subject did not voluntarily describe the goal and the parts of the problem, he/she was asked (1) What is the goal of this problem? and/or (2) What are the parts of this problem? Additional probes were not asked for each problem part because this would essentially be identical to probing the instructional materials. For example, each problem had an X axis and a Y axis (some also had a Z axis). If each subject were asked "what is an X axis," this would be tapping into the instructional knowledge structure which is what the instructional question probes were designed to do. In other words, all elements of the problem statements were elements of the instructional tape, and were therefore not probed beyond the first level described above.

Nine problems for subjects to solve were developed consisting of a 2-D or 3-D sketch of a vector with instructions to write the vector in mathematical form (see Figure 1). The problems ranged in difficulty from being relatively easy to being difficult but solvable. The problems broadly fell into three categories, easy, moderate, and difficult. Each subject did one

10

Figure 1. Sample vector problem

Given below is a graphical representation of a vector.
Write the vector in mathematical form.
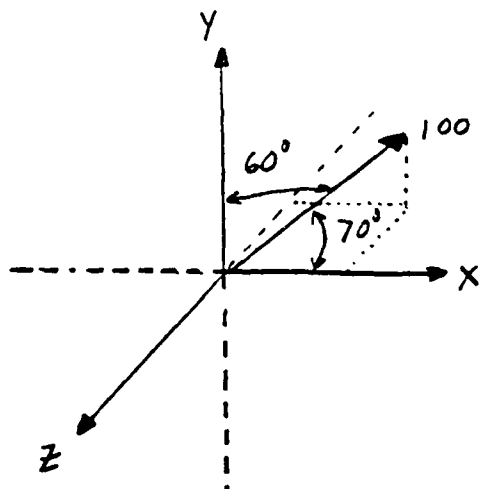
problem of each type (three total).

A post-experiment questionnaire was developed asking for subjective ratings of the videotape and problems. Questions included both a 7-point rating scale and an open-ended question on clarity of the instructional tape, completeness of the instructional tape, difficulty of the instructional tape, clarity of the problems, completeness of information given in the problems, and difficulty of the problems.

Procedure. All subjects were run individually in a large room with several tables. Subjects filled out the background questionnaire and only subjects who had previously completed at least one semester of college algebra continued with the study. Subjects were asked to watch the instructional tape carefully as they would be asked to answer questions afterwards as well as to solve several problems.

Subjects viewed the instructional videotape on a Sony 13" TV and then received a five-minute break. They were given question probes and problems to solve, the order depending on subject group. Question probe answers were tape-recorded using a portable audio cassette player. Subjects were then administered the post-experiment questionnaire and debriefed.

Results and Discussion

Instructional Tape Evaluation. Most subjects rated the clarity of the tape as being fairly high (mean = 2.1 on a -3 to +3 scale) and the tape relatively complete (mean = 2.0). When asked whether there was anything that could be added or changed to make the tape clearer or more complete, several subjects suggested spending more time on math, equations, and procedures.

Instructional Question Probes. Although the time required to answer all instructional question probes averaged about 30 minutes, subjects seemed to tolerate the process well, and found the task not too difficult. The number of "I don't know" responses ranged from two to 26 (with a mean of about 10 out of 72). Most of the questions seemed straightforward and easily understood by the subjects.

Questions were reviewed primarily to determine whether any specific questions or any types of questions were problematic for any reason. A small proportion of questions seemed to elicit several answers of the type "I don't understand" or "I'm not sure what you mean." There seemed to be two types of question where this happened more frequently: (1) questions asking for subtypes of some concept, such as "What are the types of vectors" (note that in the videotape, the subtypes were concepts discussed at some later point and the subtype relationship was only implied in the instructional materials); and (2) questions relating a concept to a visual graphical drawing in the instructional tape (such as "how is a 2-D coordinate axis represented" or "how is the positive Y axis represented"). Some subjects also seemed confused by the use of the term "properties."

Problem Question Probes. Virtually all subjects described the problem in the same manner. They stated that the goal was to write the vector in mathematical form (as explicitly given in the problem statement). In describing the graphical representation, subjects listed the axes given (X, Y, and in some cases Z), described the vector in terms of the direction it was heading, listed the magnitude given in the problem, and usually also described the vector in terms of the angle(s) given. In summary, subjects essentially simply described all visual parts of the problem and affirmed that the goal was to translate the graphical representation into a mathematical form.

Evaluation of Problems. Problem solving performance was evaluated for all nine problems (each was attempted by 8 subjects). Problems were scored by subject matter expert who has taught the material for several years. The scoring scheme was a standard instructional method; each problem was worth some total number of points (either 30 for 2-D or 50 for 3-D problems) and some number of points were deducted for each error.

The problems varied in difficulty mainly due to two factors; the number of dimensions (two vs. three), and placement of the vector relative to the axes (sometimes necessitating movement and redrawing of the vector). Although subjects did not rate the problems as unclear or incomplete, they nevertheless had difficulties with virtually all of them. Mean scores for the four 2-D problems ranged from 17.7 to 19.6 out of 30 points. Mean scores for the five 3-D problems ranged from 25.2 to 32.9 out of 50 points. For the 3-D problems, less than half of the subjects were able to completely and successfully solve any given problem.

Summary. Overall, the question probe method seemed to work well in eliciting subject knowledge of the instructional material. Only a few of the instructional question probes proved confusing or otherwise problematic for subjects.

On the other hand, the problem statement question probes seemed to be of limited value. They only elicited a verbal description of the graphical material, and any further probing would simply tap into the representation of the instructional tape.

EXPERIMENT 2

The primary purpose of Experiment 2 was to determine whether the administration of instructional question probes affected subsequent problem solving performance in the engineering mechanics domain. That is, by virtue of thinking about the instructional material and organizing answers to the various question probes, subjects might strengthen and or re-organize knowledge structures in a way beneficial to accessing that knowledge for problem solving purposes. The primary manipulation to evaluate intrusiveness of question probes was the administration of question probes before problem solving for some subjects but not others. It was hypothesized that if question probes are intrusive (presumably in a beneficial manner), problem solving scores would be higher for subjects who received the set of questions before solving the vector problems.

If the question probes turned out <u>not</u> to be intrusive, then question probe and problem solving data from the experiment could be analyzed individually for each subject who had been given question probes before problem solving. Thus, the second goal of Experiment 1 was performing a "within-subjects" evaluation of the use of conceptual structures in problem solving. The goal of this analysis would be to predict problem solving performance as some function of operations performed upon subject's knowledge structures. The knowledge structures would be operationally defined as the conceptual graph formed from answers to question probes.

The question probes developed in the EM pilot study seemed to naturally fall into two categories, basic concept or <u>WH</u> questions (what and why questions such as what is, what are the properties, what are the parts, why, etc.), and procedural or <u>HOW</u> questions (what is the procedure for X, how do you do X, what do you do when X, what is the equation for X, etc.). We decided to differentially test the intrusiveness of these two categories of questions. It seemed possible that thinking about the basic concepts and their relationships could have a different effect on problem solving than recalling specific procedures and equations used in solving actual problems. Therefore, intrusiveness was evaluated separately for the WH set of basic concept questions, the HOW set of procedural and equation questions, and BOTH sets combined.

Method

Overview. Subjects were first given a questionnaire screening for relevant mathematical education and then asked to watch a 50-minute instructional videotape on writing vectors. After a break, they were given a set of question probes and asked to solve four problems. The experimental manipulations consisted of varying the specific set of question probes and the presence or absence of question probes before problem solving. There were four experimental groups: (1) subjects who were given WH question probes before problem solving, (2) subjects who were given HOW question probes before problem solving, (3) subjects who were given both types of question probes before problem solving, and (4) a control group who were not given question probes of any type before solving the problems.

Subjects. Subjects were 60 male and female University of Idaho Introductory Psychology students who received course credit and $2/hour for participation. Subjects also received $2 for each problem solved correctly. Four subjects were replaced; three because the experimenter skipped question probe items and one because the subject gave up on problem solving after the first problem.

Materials. Materials consisted of a 50-minute instructional videotape, a subject background questionnaire, various sets of question probes, and four problems requiring subjects to model vector equations. Each of these materials are described below.

The results of the pilot study post-experiment questionnaire were used to revise the videotape one last time. This revision consisted of adding a brief summary of procedures and equations at the end. Since no new material was added, the graph form of the videotape remained that listed in Appendix D.

Based on results of the EM pilot study, any question probes that were confusing or otherwise problematic were either rewritten or eliminated. The modifications can be seen in the right column of Appendix I. Any questions deleted are shown in parentheses and any added questions or parts of questions are shown in boldface. The modified set consisted of 63 question probes. Each question probe is listed next to the relevant node-link-node in the instructional tape conceptual graph (left column).

The question probes were next categorized and divided into two types: (1) WH questions such as What is X, What are the properties of X, and Why do you X; and (2) HOW or procedural questions such as How do you X, What is the procedure for X, What happens after you X, What is the formula for X, etc. There were 41 WH questions and 22 HOW questions. A third set of question probes consisted of BOTH sets of questions combined (63 questions).

Problem solving scores from Experiment 1 were used to choose four problems that ranged in difficulty, two in 2-D and two in 3-D. One problem was modified slightly resulting in the need for one more calculation after moving the vector.

Procedure. Subjects were run individually in a large laboratory with several tables. Each subject was asked to sit at a table facing a television where the instructional tape was to be played. The subject was asked to read and sign an informed consent form describing the general purpose and procedure of the experiment.

Subjects were then asked to fill out the subject background questionnaire. All subjects who had previously had a mathematics class at the college algebra level or higher continued with the experiment. The experimenter asked subjects to watch an instructional videotape, telling them that they would be asked to solve several problems applying the information that they learned from the videotape, and that they may also be asked to answer some questions about the material.

Subjects watched the instructional tape and then took a five-minute break. They were then asked question probes or given problems to solve depending on the experimental condition. Three of the four groups (the WH, HOW, or BOTH conditions) were first given the appropriate set of question probes. The experimenter read each question out loud and subjects responded verbally, with all responses tape-recorded. After completion of the question probes, subjects were given four problems to solve, one at a time and in order of increasing difficulty. Each problem was typed and drawn at the top of a page and subjects were asked to write out in detail their answers below. After each problem solving attempt subjects were required to turn in their answers before moving to the next problem. There were no time limits for any of the problems.

Subjects in the control group were given the four problems to solve immediately after the five-minute break. All subjects were thanked and debriefed after the problem solving task.

Results

Subject problem solving scores were analyzed to determine whether question probes affected problem solving performance. Each problem was scored by an engineering subject matter expert blind to the subject's experimental condition. The scoring scheme was the same as that used in Experiment 1; the two 2-D problems were worth 30 points each and the two 3-D problems worth 50 points.

Mean problem solving scores for the four groups are shown in Table 2. Analysis of Variance revealed no significant differences between the four groups for any of the problems nor for subject's total problem solving scores ($F < 1$).

Conceptual Graph Analysis

Because there was no evidence for question probe intrusiveness, further analyses were performed for the data from subjects who received both types of question probes (WH and HOW) before problem solving. We hypothesized that subjects search the conceptual graph during problem solving by matching the problem

15

Table 2. Scores for engineering mechanics vector problems as a function of question probe condition.

Question Probe Condition	Problems				
	1	2	3	4	TOTAL
WHAT questions	25.6	25.6	44.6	40.7	136.5
HOW questions	25.4	22.3	38.3	34.2	120.3
BOTH what and how	25.8	24.6	35.8	31.5	117.7
No Question Probes	24.6	23.8	37.5	32.5	118.4
Total Possible	30	30	50	50	160

concepts onto corresponding nodes in memory. Search is carried out by using spreading activation processes to access relevant structures. Acceptance of this view leads to the prediction that the contents of the conceptual graph should be highly predictive of performance. This possibility was investigated in two ways: (1) Graphs were scored on accuracy by matching them with the instructional graph and categorizing all additional sections. These scores were used to predict combined problem solving scores; and (2) Each individual's graph was evaluated to derive a specific description of predicted problem solving activity for each of the problems.

Predicting Scores with Graph Accuracy. Question probe answers were transcribed and translated into conceptual graph form. Similar to the instructional graph, the graph nodes could be either simple concepts or more complex ideas such as propositional statements. Any node containing multiple concepts represented a more molecular network, all of which was tied to any links associated with the node.

Two judges independently scored the node/link/node segments as falling into one of five categories:

(1) Match. The node-link-node segment represented by the subject's statement was essentially the same as one of the node-link-node segments in the instructional graph.

(2) Correct New Link. The node-link-node segment contained nodes that were essentially the same as two nodes contained in the instructional graph but the segment link was not existent in the instructional graph (representing an inference linking two concepts presented in the instructional material). In addition, the link was judged to be a correct inference.

16

(3) <u>Incorrect New Link</u>. Same as (2) only the link was judged to be an incorrect inference.

(4) <u>Correct New Node</u>. The node-link-node segment contained at least one completely new concept (node) that was not contained in the instructional graph (representing an inference that drew on the subject's previous knowledge). The information in the segment was judged to be correct.

(5) <u>Incorrect New Node</u>. Same as (4) only the segment was judged to be incorrect.

The scores for the two judges were compared and any differences were resolved by a third party. The percent of agreement between the two independent scorers ranged from 79% to 96% for the 15 subjects. Frequencies were tabulated for each category. Table 3 lists the summary statistics for the frequencies in each category type. It can be seen that subjects gave answers resulting in an average of about 113 node-link-node segments (the instructional graph had 204), with a wide range from 69 to 142. A correlational analysis indicated a significant relationship between total number of segments (node-link-nodes) given in question answers with problem solving performance, $r = .49$, $p < .05$.

As indicated by the statistics in Table 3, there was also a wide variation in the make-up or composition of subjects' graphs in terms of the completeness (category 1), correctness (categories 1, 2, and 4), and amount of material imported from previously existing knowledge structures (categories 4 and 5). The average match of subject graphs to the instructional graph was about 76 links, with a wide variation among subjects. In addition, subjects added an average of 11 new links between instructional nodes and about 26 new "imported" nodes.

As a preliminary evaluation of the relationship between subject graphs and problem solving performance, the absolute scores for each of the five categories listed in Table 2 were used as predictors for subject's total problem solving scores (total of points received for all four problems). In addition, two different composite scores were analyzed; (a) total number correct was calculated by summing categories 1, 2, and 4, and (b) total number incorrect by summing categories 3 and 5.

The top row in Table 4 shows the correlation coefficients for each of the five variables. Statistical tests of the coefficients showed that all but "number of correct new links" were significantly correlated with performance. In addition, a multiple regression analysis using the five categories as predictors of performance resulted in $R = .88$ ($p < .01$) and $R^2 = .77$. However, most of the performance variance was accounted for by category 3, number of incorrect new links. In addition, the regression using the five categories was more predictive of performance than using the composite score of number correct ($r = .75$) or the composite score of number incorrect ($r = .78$).

Table 3. Summary statistics for category frequency scores
and total number of links.

	Mean	SD	Min	Max
(1) Match	75.6	24.0	35.0	112.0
(2) Correct New Link	4.3	2.4	1.0	10.0
(3) Incorrect New Link	7.1	6.2	1.0	23.0
(4) Correct New Node	17.2	9.9	6.0	43.0
(5) Incorrect New Node	8.5	8.8	0	36.0
TOTAL	112.7	21.3	69.0	142.0

The boost in prediction by adding the other categories to
category 3 suggested that the proportion of incorrect to correct
segments could be critical in accessing various parts of the
knowledge structure during problem solving. That is, almost any
model postulating search of a declarative knowledge structure
would predict that accessing and using a correct link would be
less likely if there were also many incorrect links. Accord-
ingly, the number of segments in each category were recoded as
proportions out of the total number of segments in the subject's
graph. These category proportion or percent scores were used as
predictors of performance scores in a regression analysis similar
to analysis just described. The results of the analysis were
almost identical: The percent of incorrect new links was most
highly predictive of problem solving, $r = -.88$ ($p < .01$) and
$r^2 = .77$. Combining the two "error" categories 3 and 5 into a
total percent incorrect score reduced the correlation to .83.
This indicates that the incorrect information imported onto the
graph was very peripheral to problem solving. It was the
incorrect associations among the instructional concepts that was
mostly responsible for subject errors.

Individual Performance Prediction. The previous analysis
showed a strong relationship between the accuracy of an
individual's graph and problem solving performance. However,
this relationship was only a "global" one. That is, the
compositional correctness of the conceptual graph only predicted
problem solving performance in general (i.e., total score).

Table 4. Correlation between category frequencies/proportions and performance.

	Category				
	(1) Match	(2) Correct New Link	(3) Incorrect New Link	(4) Correct New Node	(5) Incorr. New Node
Frequency	$.65^{**}$	.22	$-.82^{**}$	$.53^{*}$	$-.70^{**}$
Proportion	$.66^{**}$	.09	$-.88^{**}$	.42	$-.75^{**}$

* $\underline{p} < .05$
** $\underline{p} < .01$

If the conceptual graph represents, to some unknown degree, the knowledge structure used by subjects during problem solving, then analysis of that graph should lead to specific predictions for actual problem solving activities on an individual basis.

The goal of this analysis was thus to individually analyze each subject's conceptual graph and predict actual problem solving behavior during the course of the four problems.

Our ultimate goal in this line of research is to map declarative knowledge structures, hypothesize operations used on those structures, and predict the specific <u>sequence</u> of steps in problem solving performance. To do this, it would be necessary to have data from complete think-aloud protocols of problem solving. In the present study, subjects were simply asked to try to solve the problem and write their answers on paper. Because this could result in many "missed" steps, the sequential predictions were not attempted in this study.

For this analysis, we tested a very simple model that assumed complete operational access to the declarative knowledge structure. More specifically, the model assumed:

(1) Subjects initially learn a domain by storing information in associative knowledge structures.
(2) When subjects receive a problem to be solved, the concepts in the problem statement map onto or match concepts in the knowledge structure.
(3) Activation spreads from the nodes activated by the problem statement to related nodes.
(4) Activation will spread to associated concepts and segments of various types. For example, in the vector material, activation will spread to procedures, equations, basic concepts, etc.

(5) Any segments providing directly usable "action" knowledge (i.e., procedural sections containing "do X") will be acted upon as a step in problem solving. That is, activation spreading to areas of the graph with implementable actions will have the most direct initial effect. This, in effect, bypasses the need to derive actions from more abstract or non-action conceptual knowledge.

(6) If no directly implementable information is activated, activated conceptual knowledge may provide enough information to generate an action. That is, procedural knowledge is constructed on the basis of concept knowledge. (However, in this particular study, we assumed that the difficult nature of the material would result in _very little_ of this type of inferencing; only when very straightforward application of conceptual knowledge is possible. For example, some subjects knew that one property associated with a vector arrow pointing in the direction of the positive Z axis (that is, "out" of the page) is that the cosine θz is positive. Thus if the arrow in the problem points out of the page, they will perform the step of making the cosine θz a positive number in the unit vector equation.

To test this model, we first developed a procedural outline of the basic steps necessary to solve each problem (the outline was hierarchical where appropriate). The number of steps in the procedural lists ranged from 8 to 17 for the four problems. One experimenter then evaluated each conceptual graph to predict which steps would be successfully performed by the subject. Essentially, if the nodes activated by the problem statements were associated with correct procedures or equations for a given step, then performance of that step was predicted to be successfully executed. If procedural information was incorrect, then incorrect performance was predicted. If no procedural information was present but conceptual information was adequate to _directly_ provide a procedure (see example under assumption #6 above), then correct performance was predicted. If inadequate procedural and conceptual information was present, the prediction was an absence of that particular step.

For an initial unbiased analysis, we had the subject matter expert use the same procedural checklist and score each subject on whether their answer reflected the successful completion of that particular procedure or step. However, it should be noted that sometimes this was a difficult inference as some subjects skipped numerous steps and ended up with an answer that was some deviation from the correct one. This made determining what steps had been accomplished, and accomplished correctly, somewhat difficult and subjective.

Next, for each problem step, we compared the prediction based on the subject's graph with the SME judgment of whether the step had been accomplished. Each comparison was scored using the following scheme:

20

```
(1) + +  predicted success, actual success
(2) - -  predicted failure, actual failure
(3) + -  predicted success, actual failure
(4) - +  predicted failure, actual success
```

Mean frequencies for each category are presented in Table 5.
It can be seen that frequencies for the four problems were
relatively comparable. For all four problems, predictions based
on conceptual structures were accurate for about 87% of the steps
(categories 1 and 2 combined) with the majority being correct
prediction of successful performance on a subtask. For those
steps where the predictions were not accurate (categories 3
and 4), the errros were divided approximately equally between
predicting a success when the subject actually failed to carry
out the subtask, and predicting failure when the subject actually
did complete the subtask.

Summary and Discussion

The results of this study can be summarized as follows:
(1) The administration of question probes eliciting subjects'
"static" knowledge (as opposed to asking for new inferences) did
not significantly affect problem solving performance in either a
beneficial or deleterious manner; and (2) Individual graphs of
subject knowledge structures were strongly predictive of actual
problem solving performance. This second effect was evidenced in
two ways. First, there was a strong correlation (.88) between

Table 5. Mean frequency for four scoring categories and mean
number of correct predictions (percentages given in parentheses).

|	Problem			
	1	2	3	4
Scoring Categories				
(1) + +	5.2 (65)	8.5 (71)	8.1 (67)	11.3 (66)
(2) - -	1.7 (21)	1.9 (16)	2.4 (20)	3.5 (20)
(3) + -	.5 (06)	1.0 (08)	.9 (07)	1.2 (07)
(4) - +	.5 (06)	.5 (04)	.7 (06)	1.1 (06)
Total Number of Steps	8.0	12.0	12.0	17.0
Correct Predictions (Categ #1 and #2)	6.9 (.86)	10.4 (.87)	10.5 (.87)	14.8 (.87)

the proportional accuracy of individual graphs and the total problem solving score. The second way was that given some basic search assumptions, the information contained in each individual's conceptual graph was used to predict a majority (87%) of the steps used by subjects during problem solving.

The predictiveness of the conceptual graphs can be considered extremely high given certain mitigating factors in the study. First, the assignment of graph match and accuracy scores involved some small but definite amount of subjectivity. The propositional segments given by subjects in their answers to question probes had to be rated as being similar or different from any segment contained in the instructional graph. In addition, some of the pieces of information "imported" from subjects' previous knowledge structures were not either correct or incorrect so much as irrelevant, so these distinctions were occasionally arbitrary.

Second, and probably most important, the problem solving scenario was not set up to be conducive to step by step analysis. Subjects were only asked to write down their answers and although subjects tended to write intermediate steps, it was obvious that many steps were not written down. This made it necessary to infer what intervening steps were or were not essentially or "successfully" accomplished. This loss of direct data may have led to errors in describing subject problem solving activities, lowering prediction accuracy.

Finally, a third problem was that subjects sometimes thought of a particular concept or procedure during question probes but not problem solving, and vice versa. One possible explanation is that subjects have the information at one time but not the other. There is no obvious reason for why this difference would occur. It is more likely that subjects have the knowledge at both times but simply fail to retrieve or activate the information at one time or the other. This could possibly be due to the fact there were stimuli that activated a given concept under one circumstance but not another. That is, given an associative model of memory, the activation of a concept under both circumstances would depend on the degree of stimulus set match between the two circumstances. Unfortunately, the stimuli present during question probes were sometimes different than during problem solving. For example, Figure 2 shows the graphic part of a simple 2-D vector problem. One thing that subjects must do to write the vector in mathematical form is calculate the angle between the vector and the X axis. With the visual graphic to look at, it will be quite "obvious" to most subjects that one simply subtracts the angle from 90. However, in eliciting that knowledge during question probes, the primary stimulus consists of the question "how do you find θx given θy?" Without the graphic acting as a cue, it is not suprising that some subjects answered the question incorrectly (or failed to answer) yet did not have trouble performing the procedure during problem solving. Thus, one goal for future work should be to identify means for more completely accessing knowledge structures, especially those associated with spatial information rather than just verbal information.

Figure 2. Sample problem graphic
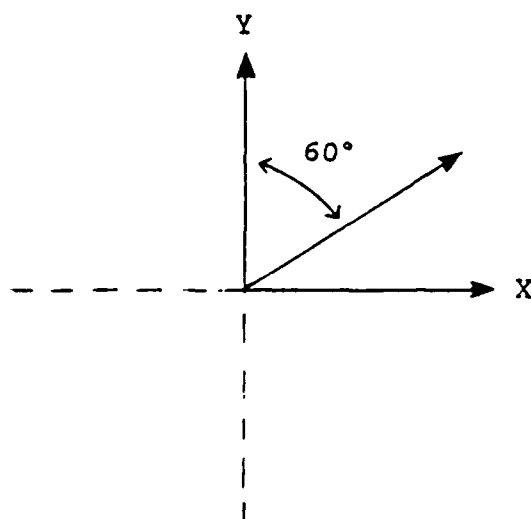
EXPERIMENT 3

The purpose of Experiment 3 was to develop and test
experimental materials, including question probes, for the domain
of video recording. The question probe group answered a
background questionnaire, saw an instructional videotape,
answered an extensive set of question probes, solved seven
problems requiring use of the VCR and peripheral equipment, and
answered a post-experiment questionnaire. A control group
received the same sequence without the administration of question
probes.

Method

Subjects. Subjects were 24 male and female University of
Idaho Introductory Psychology students participating for course
credit. Subjects were randomly assigned to either the question
probe or the control group.

Materials. A subject background questionnaire was developed
to determine experience in operating VCRs. The questionnaire
requested information on age, gender, number of times subjects
had used VCRs to perform various functions, whether they owned a
VCR, and whether they had ever done any audio-dubbing with a VCR.
To develop the final instructional tape, a preliminary tape
was recorded and translated into a conceptual graph. Based on
the graph, a new videotape was produced. The videotape consisted
of a 40 minute lecture on certain functions of hi-fi VCRs, how
they differ from mono or stereo VCRs, and a subset of the

23

controls and displays of the JVC Hi-Fi VCR.

In this study, only questions probing the instructional knowledge structure were developed and evaluated. To develop the question probes, a conceptual graph in list form was made for the final instructional tape. Appendix II lists, in the left column, a small subset of the entire graph. The graph contained 393 nodes, considerably more than the number of nodes in the engineering mechanics graph (171). The links or relationship between nodes were primarily "is a, set (subset), part, property, procedure, manner, outcome, reason, and instance."

After listing the node-link-nodes for information in the instructional videotape, question probes were written corresponding to the link types for each node. The right column of Appendix II lists question probes written for the conceptual graph segments in the left column. The probes consisted of 89 questions such as "What are the properties of the audio monitor button; What are the settings of the audio monitor button; How do you choose a source to record onto a tape on the JVC; What happens when you push the slide switches to the right?

Seven problems were developed, ranging in difficulty, asking subjects to use a JVC Hi-Fi VCR and various peripheral pieces of equipment such as a TV and auxiliary stereo-cassette player. Each problem was typed on a piece of paper along with the alloted time. Time allocations varied from four to ten minutes.

A post-experiment questionnaire was developed to aid in evaluating the experimental materials. Questions consisted of a -3 to +3 7-point rating scale and an open-ended question on each of the following; clarity of the instructional tape, conpleteness of the instructional tape, difficulty of the instructional tape, difficulty of the problems, and clarity of the problems.

Procedure. All subjects were run individually in a large room with several tables. Subjects sat at a large table with a JVC VCR facing them. They filled out the background questionnaire concerning their experience with VCRs and then were asked to watch the instructional videotape. They were told that they would be asked to answer questions afterwards as well as to solve several problems. After viewing the videotape, subjects received a five-minute break.

Subjects in the question probe group were moved to a table away from the TV and VCR equipment and verbally administered all question probes by an experimenter. Their answers were recorded on a portable cassette recorder. After another five-minute break, subjects were asked to perform seven problems using the VCR. To keep a record of problem solving, subjects were asked to verbally instruct the experimenter on the procedures to use to solve the problem. That is, the subject would tell the experimenter some procedure to perform and the experimenter would then perform that procedure on the JVC VCR in front of them on the table. Subject instructions to the experimenter were recorded on stereo cassette players. After the alloted time, the experimenter noted the settings on all equipment involved in the problem and reset the controls for the next problem.

After solving, or attempting to solve, all seven problems

subjects were given the post-experiment questionnaire, thanked, and debriefed. Subjects in the control group were run using the same procedure with the exception of the question probes.

Results and Discussion

Subject Background. The background questionnaire indicated that almost all subjects had some experience with using VCRs, mostly from either playing a tape for TV viewing or recording a TV program. However, most had not used a VCR to record from another VCR (75%), and had not used a VCR to edit or combine videotaped material from another source (87%). Only six had used a Hi-Fi VCR that they knew of and only one had ever done any audio-dubbing using a VCR.

Instructional Tape Evaluation. The instructional videotape was rated relatively high on clarity (modal response was 2 on -3 to +3 scale) and completeness (modal response was 2). The tape was rated as being moderately difficult to understand, $\underline{x}$ = -.05 on a 7-point scale with -3 being very easy and +3 being very difficult. Several subjects commented that there was a lot of material to absorb in 40 minutes. This is interesting considering that a fairly high proportion of the instructional material covered basics of VCRs that the subject would be expected to know from previous experience. Given the ratings and comments on the questionnaires, we decided to modify the instructional material and produce a new tape for Experiment 4.

Question Probes. Although they rated the instructional videotape as being relatively difficult to understand (and remember), subjects were able to give some kind of answer to almost all of the 89 questions. Three subjects answered all of the questions, and the mean number of "I don't know" responses for the other nine subjects was 8. In addition, there were virtually no questions where subjects indicated that they did not understand what was being asked of them. The question probes chosen to elicit the VCR knowledge seemed successful and appropriate and no further analysis was conducted.

Problems. Subjects seemed to readily accept the constraint of asking the experimenter to perform the tasks to solve a given problem. To score subject performance, each problem was broken down into the correct procedures necessary to complete the task. The audio recordings and post-problem instrument settings were then used to tally the procedures correctly or incorrectly performed by subjects. The score on a given problem was the number of correct procedures out of the total number possible. The problems varied widely in terms of the number of subjects who were able to successfully complete the task (ranging from 0 subjects to 22 out of 24). Based on problem clarity ratings and subject performance, four problems were chosen and slightly revised for Experiment 4. The problems were expected to provide a wide range of difficulty; the number of subjects successfully completing these four problems were 18, 14, 9, and 3.

EXPERIMENT 4

Experiment 4 was a conceptual replication of Experiment 2; the primary goal was to test the intrusiveness of question probes in the VCR domain. If the question probes did not affect subsequent performance, the conceptual graphs would be analyzed and used to predict problem solving performance on an individual basis. The design of Experiment 4 was similar to that for Experiment 3; one experimental group received question probes while the control group did not. However, subjects solved four problems rather than seven and no post-experimental evaluation questionnaire was given.

Method

Subjects. Subjects were 20 University of Idaho Introductory Psychology students participating for course credit and $2/hour, and 10 University of Idaho students participating for $5/hour (randomly split between the two experimental conditions). In addition, subjects received $1 for each problem solved correctly. All subjects were randomly assigned to either the quest or the control group.

Materials. The subject background questionnaire used for Experiment 3 was used without revision for Experiment 4. Results from Experiment 3 were used to make minor revisions to the instructional videotape. Most of the revisions consisted of cutting out or reorganizing material. A new videotape lasting 35 minutes was produced.

The conceptual graph was redeveloped as a result of producing a new videotape. The conceptual graph was then used to finalize the question probes to be used in Experiment 4. During the process of finalizing the conceptual graph, it was decided that the question probes had worked well for Experiment 3, but there were still segments of the graph that were not being probed adequately. As a result, 43 new questions were added to the set resulting in a total of 132 question probes. The questions all generally fell into one of eight types; PROPERTY (35 questions), HOW (33), OUTCOME (30), WHAT IS (11), SET such as "what are the types of...(10), WHY (7), HOW DO YOU KNOW (5), and INSTANCE (1). Unlike the vector study, there did not seem to be any clear and simple break in the type of questions, therefore no differentiation into a small number of categories (such as basic concepts vs. procedures) was attempted.

Four problems of varying difficulty were selected from the seven used in Experiment 3. To facilitate later discussion of conceptual graphs and problem solving, each will be briefly described along with a list of the most critical knowledge structure "areas" necessary for completing the task.

Question 1: Set up the system so that you are watching channel 10 on the TV through the VCR tuner.
   Required: (1) use of TV/Video toggle switch

(2) Video mode is used to bring a signal into the
VCR tuner and sent on to the TV

Question 2: Record channel 10 on tape, and at the same time watch
channel 6 on the TV.
  Required: (1) use of TV/Video toggle switch
            (2) TV mode is used to bring one signal into VCR
                and also have all signals bypass VCR to go to TV
            (3) all functions associated with recording a tape

Question 3: Given the recorded tape in front of you, instruct the
experimenter to perform whatever steps are necessary
for you to determine what this tape would look and
sound like if played on a mono VCR.
  Required: (1) how mono VCRs play tapes
            (2) how Hi-Fi VCRs play tapes
            (3) recordings on different sections of the tape
            (4) properties of audio monitor button (allow to
                listen to recording)
            (5) setting of audio monitor button ("normal")

Question 4: The audio cassette player is connected to the auxiliary
input of the VCR. Make and play a tape with the
following characteristics:
            - The tape, when played, plays TV video and audio.
            - After a short period of time, cassette music
              gradually fades in and plays for about one minute
              and then fades back out.
  Required: (1) use of Source Select button to receive input
                from both TV and auxiliary source
            (2) section where TV audio signal is put on tape
            (3) section where auxiliary audio signal is put on tape
            (4) audio monitor button (used to listen to various
                parts of the tape, or different signals)
            (5) all functions associated with actually
                recording a tape (tape in, hit Rec/1tr, etc.)
            (6) use of record level slide switches to bring
                music in and out
            (7) use of ALC button to allow manual control of
                hi-fi sound recording (music component).

The problems were each typed at the top of a page along with the
time allowed for problem solving.

   Procedure. The procedure used in Experiment 4 was identical
to that used for Experiment 3. Half of the subjects received
question probes before problem solving while the other half did
not. Subjects solved the problems one at a time, instructing the
experimenter how to perform the task on a VCR in front of them.
Subject answers were tape-recorded using a portable cassette tape
recorder.

Results

Subject Background. As in Experiment 3, subjects exhibited some familiarity with the use of standard VCRs. Most of this background was limited to using a VCR to record a TV program or play a tape on a TV set.

Question Probe Intrusiveness. Each of the four problems were scored by tallying the number of procedural steps performed correctly. Problems 1 through 4 had two, five, two, and nine necessary steps, respectively. Because of the low number of points possible for problems 1 and 3, intrusiveness was evaluated by combining points for all four problems for each subject, resulting in a total possible of 18. Mean scores were 12.8 for the group who received question probes and 12.3 for the control group, a difference that was not statistically significant ($t < 1$).

Conceptual Graph Analysis

As in Experiment 2, the lack of intrusiveness was taken as a sign that individual analyses could safely be performed. Answers to question probes were segmented and mapped into fifteen individual conceptual structures. Each structure was mapped onto the original instructional graph so that differences could be easily identified.

Predicting Scores with Graph Accuracy. Each node-link-node was rated by two experimenters as belonging to one of five categories; match with instructional graph, correct new link, incorrect new link, correct new node and incorrect new node. Table 6 shows the summary statistics for the five categories.

Table 6. Summary statistics for category frequency scores and total number of links.

	Mean	SD	Min	Max
(1) Match	140.7	29.6	97.0	199.0
(2) Correct New Link	19.5	5.7	9.0	29.0
(3) Incorrect New Link	2.7	2.4	0	7.0
(4) Correct New Node	53.5	14.9	33.0	78.0
(5) Incorrect New Node	7.5	6.2	0	20.0
TOTAL	223.9	29.2	178.0	283.0

Several things are noticeable in these statistics for the conceptual structures. First, the total number of segments obtained from the question probe data was quite high, ranging from 178 to 283 with an average of 224. This was twice as many as the number obtained in the engineering mechanics study (refer to Table 3). However, the graphs were still fairly incomplete, showing an average of only 224 links as compared to the 439 links in the instructional graph.

In addition, the proportional frequencies of the categories were quite different from the proportions found in the engineering mechanics study. As in the vector graphs, the majority of sections were correct matches, however, the number of incorrect new links was extremely small, averaging only 2.7 out of 223 links. Also, the proportion of correct new nodes was large and the proportion of incorrect new nodes was relatively small. Overall then, the graphs can be described as large and relatively accurate with few explicit errors. However, there were still many links in the instructional graph that were not evidenced in subject graphs.

Correlational analyses similar to those conducted in Experiment 2 were performed for the category frequencies and also for category proportions. These variables were used as predictors for total problem solving scores. Table 7 indicates the Pearson r values for the absolute frequency values and category proportions. However, caution should be exercised in interpreting the data because several categories, notably 3 and 5 had values very restricted in range. The data indicate that several variables were significantly correlated with performance; number of matched links, number of incorrect new links, and both number and proportion of incorrect new nodes. However, the r values are generally lower than those obtained in the vector

Table 7. Correlation between category frequency/proportion and performance.

	Category				
	(1) Match	(2) Correct New Link	(3) Incorrect New Link	(4) Correct New Node	(5) Incorrect New Node
Frequency	.64**	.23	-.54*	.28	-.59*
Proportion	.40	.46	-.51	.04	-.60*

* $p < .05$
** $p < .01$

study (see Table 4). A multivariate analysis using all frequencies as predictors resulted in an $R = .82$ ($p < .05$). The R^2 value indicates that 67% of the variance in problem solving scores can be accounted for by the composition of subject graphs in terms of the five categories.

Individual Performance Prediction. The fact that subjects verbally described the sequence of steps to be taken in problem solving provided a good opportunity to test the predictiveness of the conceptual graphs. For the initial analysis, a very simple model of information processing was assumed. Predominantly, it was assumed that the structure used by subjects during problem solving was similar to the structure experimentally obtained with question probes and that information in the structure was accessed through spreading activation in a relatively complete fashion (see detailed list of assumptions on page 20).

To develop the predictions, for each step necessary to perform the task, relevant areas of the knowledge structures were identified. If a subject's graph contained the minimal necessary information, successful performance of the step was predicted. If incorrect information was contained in the relevant section, a specific error was predicted. For example, some subjects identified "video mode" as having the property "used for recording tapes." When asked to perform a task requiring recording of a tape, subjects would conclude that the VCR must be in video mode. This knowledge would incorrectly guide the subject to set the VCR on video mode for problem #2. In this case, an error would be predicted as well as predicting an absence of the correct step, switching to TV mode.

Predictions were developed for all subjects on all four problems with the exception of three instances where no attempt whatsoever was made to solve the problem. Predictions for the remaining 261 procedural steps were compared with subject protocols. However, graph predictions were compared only to the first sequence of steps. Any additional steps taken after they had clearly failed were ignored. The reason for this is that subjects often were able to try a procedure and gain feedback (e.g., it didn't work). Failure of a procedure often led to almost random trial and error performance. This secondary trial and error performance was not included in this analysis.

The comparisons between predictions and actual performance fell into one of four categories; (+ +) for accurate prediction of the step and actual performance, (- -) for predicted omission and step actually omitted, (+ -) for predicted step but actual omission, and (- +) for predicted omission but actual correct performance of the step. The mean frequencies for the four categories are given in Table 8.

Data in Table 8 indicate that the great majority of the steps needed to perform the tasks were correctly predicted by the graphs. Of the 261 steps, 93% (all but 18) were correctly predicted; 67% were correctly predicted successful procedures and 26% were correctly predicted omissions. Looking at the data another way, of the 183 procedures correctly executed by

Table 8. Mean frequencies for four scoring categories and mean
number of correct predictions (percentages given in parentheses).

	Problem			
	1	2	3	4
Scoring Categories				
(1) + +	1.94 (97)	3.71 (74)	1.46 (73)	4.87 (54)
(2) - -	0 (00)	1.07 (22)	.31 (15)	3.33 (37)
(3) + -	.06 (03)	.21 (04)	0 (00)	.40 (04)
(4) - +	0 (00)	0 (00)	.23 (11)	.40 (04)
Total Number of Steps	2.0	5.0	2.0	9.0
Correct Predictions (Categ. 1 and 2)	1.94 (97)	4.78 (96)	1.77 (88)	8.20 (91)

subjects, 95% were correctly predicted. Of the 79 procedures
"missed" by subjects, 87% were correctly predicted omissions.
For the omissions, the predictions fell into two categories, a
prediction of omission only, and a prediction of omission
accompanied by an incorrect procedure performed in its place. In
the later case, all but one of the 19 predicted incorrect
procedures were actually performed.

Analysis of Mispredictions. The mispredictions could be
accounted for in several ways. Some of these have been explored
at this point. First, it seemed possible that some subjects in
particular might simply not have used existing knowledge during
problem solving. Or conversely, they might not have verbalized
knowledge during question probes but still used the knowledge
during problem solving. This would imply that certain subjects
would account for the majority of the mispredictions.
However, analysis of the pattern of misprediction showed that the
errors were fairly evenly spread across 12 of the 15 subjects.
This indicates that a more general phencmenon was responsible for
the result.
The cause of prediction errors could essentially be founded
in one of two sources; either the conceptual graphs did not
actually reflect knowledge held by subjects and/or the
operational assumptions (such as complete structure access) were
not correct. To gain insight into these possibilities, the
nature of the errors were described and categorized. The errors
fell into a relatively small number of categories:

31

(1) Information was in the graph, but the information was not accessed and used during problem solving.

2 subjects: said to push "play" instead of record yet graphs clearly showed they knew to press "record"

2 subjects: said to use AUX setting to get signal from auxiliary source, yet they knew that SC brought in TV and aux

1 subject: knew about the ALC switch but failed to specify setting

3 subjects: knew functions of TV mode (or video mode) but failed to specify setting

2 subjects: simply forgot to set TV channel or turn cassette player on

(2) Information was not <u>directly</u> in the graph, but subjects performed the correct task anyway.

3 subjects: couldn't specify the function of the "normal" setting for the audio monitor button, yet were able to make a small inference; to listen to a mono tape, mono is normal, so setting should be on "normal"

2 subjects: we predicted that lack of knowledge about earlier necessary procedures would prevent them from using the slide switches to increase sound, they simply proceeded without performing the previous necessary steps

2 subjects: exhibited no knowledge about ALC switch when probed but correctly turned it off to set recording level

1 subject: exhibited no knowledge about SC position yet correctly specified SC setting

Several things are noticeable about the two categories above. Regarding the first category, out of 261 steps, there were only 10 instances where the search process used during problem solving failed to access information indicated by question probes. This would imply that although not completely correct, a "complete access" search model seems to be relatively accurate in this particular problem solving scenario. In addition, we found that the information that was not accessed was either quite "distant" from the nodes activated by the problem (e.g. turning on the cassette player), or close, but an even closer node provided a wrong answer. For example, in the statement for problem #4, the word "play" was used frequently and the word "record" never used, although subjects had to record a tape to

32

perform the task. Several subjects told the experimenter to push play instead of record. One of them corrected himself but the other two ended up in the first category listed above.

It can be seen from some of the mispredictions in category (1) how representation of the problem statement can lead to an incorrect solution. For example, before problem solving, many subjects had a representation something like the simplified one shown in Figure 3. In order to record TV and cassette music in problem #4, successful subjects thought "TV and auxiliary input" and this was associated with the SC position which they specified as a setting. However, even given the same conceptual structure, other subjects focused on the "auxiliary input" component and this led them to use AUX, the wrong Source Select setting.

The errors under category (2) also provide some insight into the problem solving process. The first item under this category revealed that three subjects were able to make a <u>very simple</u> inference based on incomplete information. In the question probe test, they didn't reveal any knowledge about the audio monitor setting of "normal." However, when they searched for a way to "listen to the mono part of the tape" during problem solving,

Figure 3. Conceptual graph segment for Source Select Button.
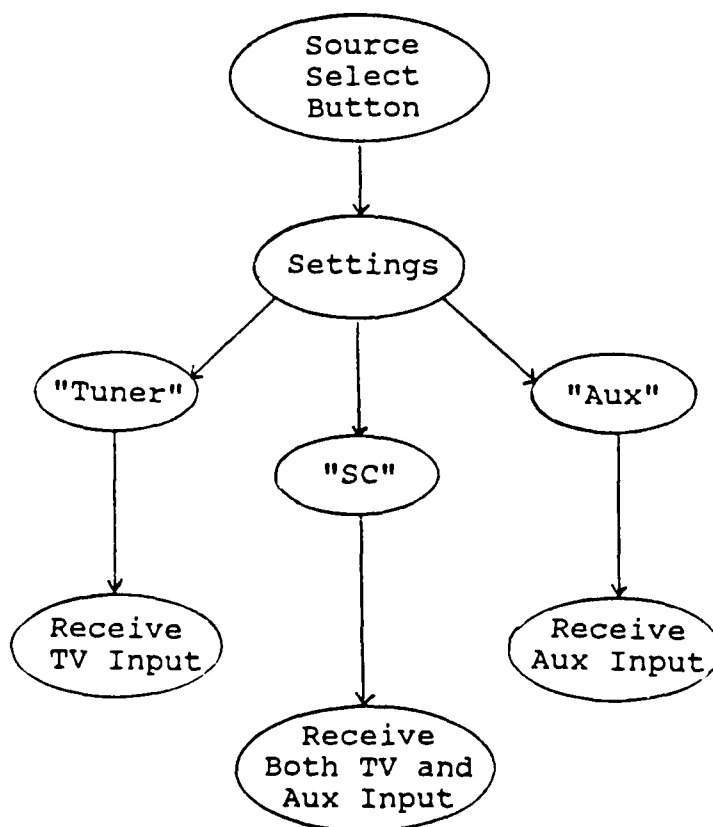
they were able to infer the answer from the following pieces of information:

1. The Audio Monitor Switch lets you listen to a tape
   (knowledge contained in their graph)
2. The Audio Monitor Switch has a setting called "normal"
   (they could see this on the system during problem solving)
3. "mono" is "normal"
   (knowledge contained in their graph)

Therefore, put the Audio Monitor Switch on "normal"

Thus, the action was missed by predictions because subjects used inferencing mechanisms. However, this process was very rare.
    The next two subjects performed a task for which they had the knowledge, however we predicted they would never get that far due to other missing structures. And, in fact, they did not _correctly_ get that far but they performed the procedures anyway.
    Finally, there were three procedures executed by subjects for which there was no underlying knowledge in the conceptual graph. This is the only potentially problematic finding of all of the prediction errors. It would seem that these subjects had some knowledge about the ALC switch (or SC position), and for whatever reason, it was not verbalized during question probes. As we stated in discussing this phenomenon in the engineering mechanics study, it seems most likely that the stimuli cueing the knowledge in the problem solving situation were simply not there during question probes. However, it should be pointed out that subjects viewed the equipment during problem solving and did not view the equipment during question probes. Given this difference, it is actually surprising that there were not more instances of this differential activation of concepts.

Summary and Discussion

    To summarize, similar to the results obtained in Experiment 2, administration of question probes did not have any statistically significant effect on subjects' problem solving performance. For that reason, individual conceptual graphs were developed on the basis of question probe answers and used to predict problem solving behavior.
    The graphs for the vector domain tended to be large, averaging 224 node-link-node segments, and were composed of mostly accurate information. The graphs clearly had previous knowledge about the use of VCRs integrated into the structure, as well as a few misconceptions that had not been eliminated by the "correct" material in the instructional tape. While subjects had relatively accurate graphs, they still represented less than half the material in the instructional tape. The portions of subject graphs detailing familiar topics such as playing tapes on VCRs were relatively complete. However the areas dealing with material new to subjects, such as the functions and settings of the ALC button, were often sparse or even completely blank.

34

The conceptual graphs were scored on variables related to old and new, correct and incorrect nodes and links. These graph composition scores were able to predict total problem solving scores relatively well (R = .88). Finally, each subject's graph was used to predict whether the subject would perform or fail to perform each of the procedural steps necessary to solve a given problem. In comparing the predictions, with actual performance, the predictions matched performance on 93% of the steps. That is, only 18 out of 261 predictions failed to match performance. These errors were composed of both hits and misses, and were outlined in the previous section.

Finally, the graphs were interesting in that subject preconceptions could be seen imbedded in the structures, even given the presence of contradictory information in the instructional tape. For example, several subjects thought that TV mode was for watching TV and Video mode was for recording tapes. Relative to the total number of node-link-nodes, these were a minor portion of the whole graph. However, the TV mode and Video mode were central concepts in the graph and thus a misconception regarding their function would significantly affect performance.

Experiment 5

The goal of Experiment 5 was to validate the conceptual graphs, that is, assess the degree to which they reflect subjects' internal associative structures. Recall clustering was used as an index of cognitive structure. Previous research has indicated that in free recall, subjects cluster or group words together on the basis of associations in memory (Bousfield, 1953). Accordingly, subjects in Experiment 5 watched an instructional tape, were asked to recall the material, and were then administered question probes. Conceptual graphs were developed from the question probe answers and the order of propositions in free recall protocols was compared with the structure of the conceptual graphs. It was hypothesized that if the graphs reflected the underlying associative structures, the sequential order of the free recall propositions should appear as clusters on the graphs.

Method

Subjects. Subjects were 12 University of Idaho Introductory Psychology students participating for course credit and $2 per hour.

Materials. Materials consisted of the VCR instructional videotape and question probes used in Experiment 4.

Procedure. Subjects were run individually in a large laboratory room with several tables (same conditions as in Experiment 4). Each subject watched the instructional tape then received a five minute break. They were then asked to tell the experimenter everything that they could remember about the

35

information in the tape. Answers to the free recall test were recorded using a portable cassette tape recorder. After another five minute break, subjects were administered the VCR question probes. Subjects were then thanked and debriefed.

Results and Discussion

Question probe answers were graphed into conceptual structures in a manner similar to Experiment 4. The free recall protocols were transcribed and broken into propositional statements. Each proposition in the recall protocol was then sequentially numbered on the corresponding node/link/node of the conceptual graph. If it was not already on the graph, it was added, numbered, and color coded as new. This resulted in a visual display showing the sequential ordering of free recall propositions, as well as the degree of match between information in the free recall protocols and conceptual graphs.

As expected, the information given in free recall was very limited relative to the conceptual structures. Recall protocols averaged about 40 statements as compared to the average of 223 node/link/node segments in the conceptual graphs.

To compare clustering in the recall protocols and conceptual graphs, it would be possible to use traditional conditional probabilities of the nodes and compare the two for conceptual graphs and recall protocols. However, given that there were over 400 nodes in all of the conceptual graphs and most of these were never verbalized in free recall, a more efficient method was chosen. To quantitatively analyze the clustering of the free recall protocols, a scoring scheme was developed representing the number of links on the conceptual graph that were "jumped" or skipped to move from one recall proposition to the next. Figure 4 shows an abstracted version of a conceptual graph segment. If

Figure 4. Abstracted conceptual graph
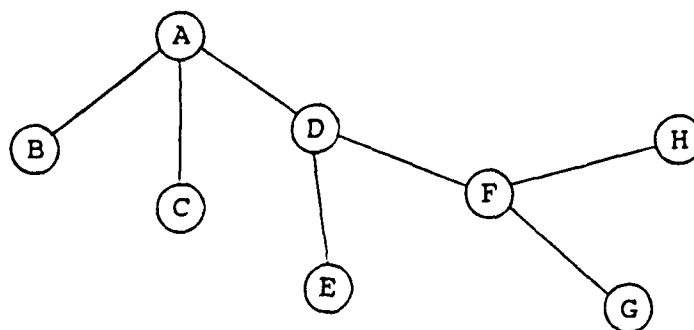
the conceptual graph reflects an underlying structure, then we would expect the free recall propositions to have a structure such as A-B, A-C, A-D, D-E, F-G, F-H, etc. The validity of the graphs would be questionable were the free recall protocols to be ordered more "randomly" such as D-E, F-H, A-B, etc.

The scoring scheme captures these differences between clustering on the conceptual graph and a more random movement around the graph. Each free recall statement was scored on the number of links that were jumped or traversed from the previous proposition. Each subject therefore had a running score reflecting the association of the recall statements in relationship to the conceptual graph. The first example described above would result in the following set of scores:

Recall Statement	
A-B	
A-C	0
A-D	0
D-E	0
F-G	1
F-H	0

Therefore, to the extent that subject's running scores are characterized by groups of 0s and 1s, punctuated by less frequent larger numbers, the conceptual graph is reflective of the underlying cognitive structure. Appendix III lists the running scores for all 12 subjects. Inspection of these data makes it clear that order of free recall was strongly clustered in the same manner as the conceptual graph organization. Most sequences consisted of small clusters of 0s separated by 1s, 2s, and occasionally higher numbers.

The free recall data indicated that there was an average of two or three concepts or nodes per subject that were not verbalized during question probes (out of about 40-50 total). More frequently, there were sometimes links in free recall that did not exist in the graphs. That is, two concepts were discussed during question probes but not with a directly specified relationship between them as in the free recall. This occurred for an average of about 9-10% of the subject recall statements.

Figure 5 summarizes the clustering data in Appendix III by giving the frequency of links jumped in going from one recall statement to the next. The data reflect numbers only for recall statements that matched the conceptual graphs; instances of new nodes and new links, as described above, were deleted. The last column, skip, represents instances where the subject went to a new graph segment not connected by any links to the previous one. The data in Figure 5 confirm that most of the recall statements for a given subject were strongly clustered in the same manner as the segments in his/her conceptual graph.

To summarize, there were two important results in Experiment 5. First, the free recall protocols were strongly related to content and organization of the conceptual graphs developed using

37

Figure 5. Mean frequency for number of links jumped to get to new recall proposition.
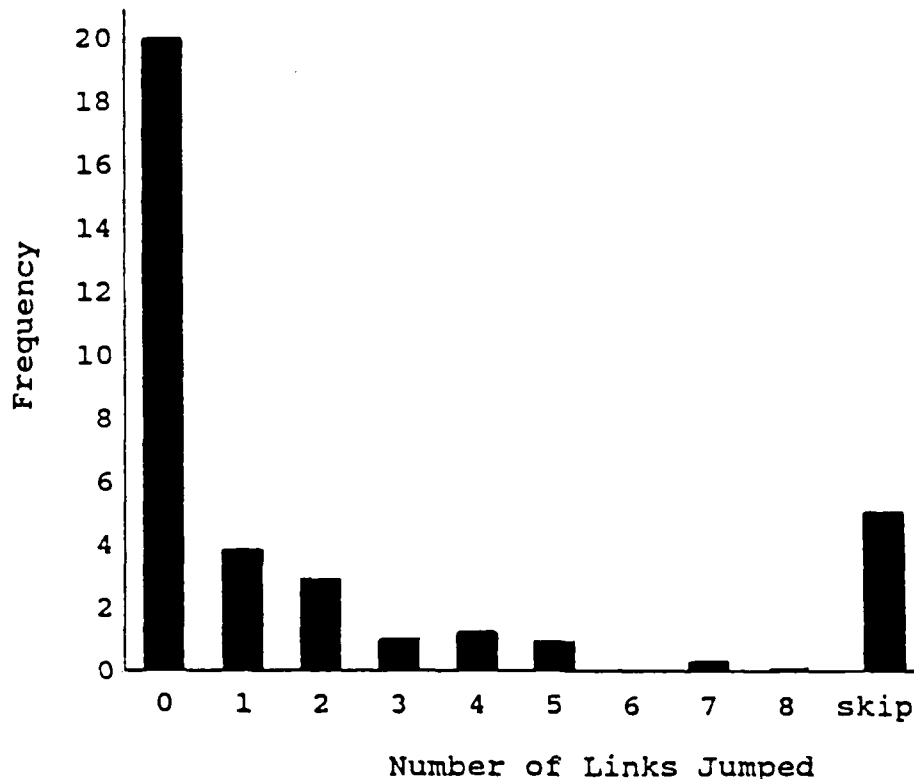
question probes. Second, subjects sometimes used simple concepts such as"button" as a way to access information for the free recall task. This association between nodes containing a common concept such as "record button" and "display button" was often not captured in the conceptual graphs. There are several possible explanations for this effect. The two most likely are that (1) subjects used a visual image of the display panel to "read off" the control buttons and display features, and/or (2) free recall included more "surface" feature associations as a recall strategy whereas the instructional material and therefore the probes concentrated more on functional associations. The possibility of associations in the conceptual structure based on surface characteristics or visual images should be addressed before more extensive modeling is done with the graphs.

General Discussion

To summarize, Experiments 1 and 3 resulted in the successful development of question probes to elicit knowledge structures in the domains of engineering mechanics and video recording, respectively. In the studies reported here, the probes were designed to be administered after subjects watched instructional videotapes. However, they can just as easily be used without prior instruction to assess knowledge of any topic. The types or

categories of question probes used for both domains were quite similar. They tended to be predominantly of the types; "What is, What are the types, What are the properties of, Why, How, and What is the result?"

Since question probes would ideally be used to assess knowledge structures just prior to problem solving, it was important to determine the extent to which they affect subsequent problem solving activity. In Experiments 2 (engineering mechanics) and 4 (video recording), we compared problem solving data for groups with and without prior question probes. We found no evidence that use of the probes changed subjects' problem solving behavior in either domain. One likely reason is that we did not ask subjects to make any inferences. That is, the probes only asked subjects about concepts and concept relationships that were previously presented in instructional material. This would therefore be unlikely to encourage subjects to think about the material in any new or fundamentally different ways.

Given that the probes were not intrusive, we used question probe answers from subjects in Experiments 2 and 4 to construct individual conceptual graphs. This graph was interpreted as being an externalized representation of the subject's knowledge structure (the "correctness" or validity of this externalization will be discussed below). Analyses showed that graphs in the engineering mechanics domain consisted predominantly of concepts that were given in the instructional tape with a relatively large number of incorrect links (averaging about 14% of the graph). There was not a great deal of information "imported" from subjects' previous schemas, presumably because most of our subjects knew very little about the topic prior to the experiment. By contrast, the graphs for the VCR domain were much larger, had very few errors, and had quite a bit of material imported from previous knowledge structures dealing with the use of VCRs. They also showed large gaps and sketchy structures in areas where the information was new from the instructional tape.

The graphs were used to determine the role of knowledge structures in domain-specific problem solving. We tested a very simple model of problem solving, assuming that a simple search mechanism such as spreading activation started at nodes activated by problem concepts. The search activated any associated procedures, basic concepts, etc. Procedural knowledge would specify actions; in the absence of procedural specifications, basic concept knowledge would be searched for directly applicable information. Little inferencing or "reasoning" was assumed. This model was tested in two ways. First, the content of each graph was numerically described by categorizing each node/link/node in terms of both accuracy and "match" with the instructional graph. These scores were predictive of total problem solving scores in both engineering mechanics ($R^2 = .88$) and video recording ($R^2 = .82$).

As a better test of the model, we used the graph content to predict the specific problem solving activities of each subject. The problems were first defined in terms of procedures necessary for successful completion of the problem. Each graph was used to predict whether that subject would or would not perform a

particular procedure. These predictions were accurate for 87% of the procedures in the engineering mechanics study (Experiment 3) and for 93% of the procedures in the video recording study (Experiment 4). In addition, for the VCR predictions of "failure to perform" a given procedure, all but one were accompanied by a correct prediction of the activity that would occur (either a simple failure to perform the step altogether or an incorrect procedure specified).[1] In summary, a model assuming relatively complete access to the knowledge structure content was able to successfully predict the specific problem solving performance of subjects at a subtask level.

To evaluate the degree to which the conceptual graphs reflected underlying cognitive structures, subjects in Experiment 5 were given both a free recall test and question probes. It was expected that if the graphs accurately reflected underlying cognitive structures, each subject's recall order would be clustered by area on his/her graph. Data analysis showed that the recall order was highly clustered by area on the graphs, with subjects giving about three node/link/node propositions in one area and then skipping to a nearby segment and repeating the process. Finally, the recall protocols only reflected about 19% of what subjects knew about the topic as measured by the graphs, indicating that question probes are much more effective in eliciting complete knowledge structures than free recall techniques.

Differences in Predictive Accuracy

The two domains differed somewhat in the predictive accuracy of the graphs. For the analysis using "match" to the instructional graph and accuracy (such as number of incorrect links) to predict total problem solving scores, the engineering mechanics scores were predicted better by total number of incorrect links, while the VCR scores were better predicted by number of segments matching the instructional graph. This is most likely due to the fact that in the vector graphs, subjects had most of the concepts there, but had many incorrect or crossed associations resulting in errors. The VCR graphs had very few errors but many areas were essentially missing. The most prevalent or influential characteristics of the graphs were simply reflected by different scoring categories.

The specific subtask predictions were slightly more accurate in the VCR study (Experiment 4) than in the engineering mechanics study (Experiment 3). There are probably at least three reasons. First, the graphs in the engineering mechanics study were characterized by many incorrect or "crossed" links: Subjects stored the concepts given in the instructional tape but there seemed to be many incorrect or "fuzzy" associations. In fact, it could easily be the case that there were many weak associations

---

[1] There was not sufficient problem solving subtask data for this analysis to be carried out in the engineering mechanics domain.

between several of the concepts, and that some of these were exhibited during question probes but others were activated during problem solving. Even if the multiple association feature were captured in the graphs (which it sometimes was), it would be difficult to predict which path would be taken during problem solving.

That fuzzy error-proneness of the engineering mechanics graphs can be contrasted with the large, relatively well-structured and error free nature of the VCR graphs. For the most part, what was there was there, and what was not was not. In fact, most subject errors were predicted not from incorrect or crossed links but simply absence of the information necessary to solve the problem.

Finally, as noted in the Experiment 2 discussion section, the problem solving data for the vector problems did not always contain intermediate procedural steps and the scorer had to infer which steps had been accomplished and which had not. When subject answers were incorrect, it was sometimes difficult to determine which steps had been performed, and performed correctly. For this reason in particular, the predictive numbers for the engineering mechanics domain can probably be considered con-servative.

Implications for a Model of Problem Solving

Central Role of Structure Content. Previous research has been able to indirectly show the relative importance of domain-specific knowledge structures in problem solving (Chi & Glaser, 1985). Most of this work has involved the comparison of novices and experts on tasks such as categorizing problems (Chi, Feltovich, & Glaser, 1981; de Jong & Ferguson-Hessler, 1986; Hinsley, Hayes, & Simon, 1978; Silver, 1981), performing memory tasks (Chase & Simon, 1973; Chi, 1978; Chiesi, Spilich, & Voss; 1979; Schneider, Korkel, & Weinert, 1989), tests of text comprehension (Schneider et al., 1989; Spilich, Vesonder, Chiesi, & Voss, 1979), and actual problem solving behavior (Lesgold, Rubinson, Feltovich, Glaser, Klopfer, & Wang, 1988). These studies and others have indicated the importance of domain-specific knowledge in problem solving. In fact Schneider et al. (1989) recently showed that domain-specific knowledge and not general aptitude accounted for differences in children's text comprehension and recall.

Consistent with this work, the current research indicates that not only is domain-specific knowledge important, it may account for almost all of problem solving behavior under some circumstances. Subjects in our studies exhibited almost no significant inferential reasoning from "basic" knowledge of the underlying concepts in a given domain. For the type of instruction and problems used in these studies, subjects used relatively simple search procedures to access relevant pieces of information. If the information was not already there, ready to be pieced together as required by the problem, subjects floundered.

It is stressed here that when subjects failed to find

directly relevant and useful knowledge in the conceptual structure, rather than operating upon the structure to come up with potentially useful inferences, their behavior became very trial-and-error oriented. This lack of substantial inferencing is reminiscent of the difficulties students have in transferring to problems different from the ones to which they are accustomed (e.g., Weisberg, DiCamillo, & Phillips, 1978). A reliance on direct memory search processes coupled with minimal use of inferencing mechanisms can account for both phenomena. This hypothesis is further supported by recent findings in educational research such as the fact that learning of "how-to-do-it" information is more helpful to students than "how-it-works" information (Lesgold, 1988; Pirolli & Anderson, 1984).

In addition, recent work on analogical problem transfer suggests that transfer based on abstract concepts is especially difficult and is a characteristic which distinguishes experts from novices (Adelson, 1984; Novick, 1988; and Schoenfeld & Herrmann, 1982). It may be the case that abstract knowledge is built up over time as part of the expert's knowledge structure, and both experts and novices still solve problems primarily through associative memory search processes. This hypothesis is consistent with Lesgold's (1988) summary of results obtained in some of his recent work with colleagues:

> "high- and low-ability people do not differ either in their knowledge of weak problem-solving methods or in their general knowledge of electrical and electronics principles. The differences lie exclusively in strategies and tactics that are specific to the kinds of troubleshooting they must do and in declarative knowledge relating to the specific levels of components about which they must make decisions." (p. 204)

Reasoning. What, then, is the role of reasoning in this view of problem solving as search of associative memory? Although speculation at this point, our data indicate that probably the first method of choice in problem solving is simple activation of directly applicable knowledge structures. Remembered information will be a combination of static knowledge of facts, procedural knowledge (either automated or otherwise), and specific instances. This is more likely an inherent function of our memory system rather than a deliberate "choice" taken by subjects. Certain weak methods of problem solving, such as means-end or subgoal analysis may be used in conjunction with the search processes if the problem is difficult or has several components.

In the absence of directly applicable knowledge, subjects will then move to the use of other strategies, not the least common of which is trial-and-error. Traditional puzzle problems, such as the Tower of Hanoi or the eight-tile puzzle, bypass the use of domain-specific knowledge and therefore automatically put subjects into this second mode. At this point, reasoning or substantial creative inference is a difficult process to which people are unaccustomed and rarely resort. This view is

consistent with evidence that direct experiential knowledge can substantially facilitate a pure reasoning problem such as Wason's abstract selection task (Evans, 1982).

Approximations to Internal Knowledge Structures

One question which naturally arises due to the nature of this research is whether the conceptual structures based on question probe answers bear any resemblance to the actual associative structures of the subjects. Experiment 5 provided strong evidence that the derived conceptual structures have a strong resemblance to internal structures because recall clustering was highly related to the organization of the graphs.

While there was a very small number of concepts in free recall protocols that were not contained in the conceptual graphs, there was a signficant number of links in the recall protocols not found in the conceptual graphs. These links were usually _type_ or category links of a very simply variety. For example, during recall, subjects would sequentially list a number of buttons or display characteristics. There are probably several reasons for this type of association being missing from the graphs. The first and most likely reason is that the question probes did not address the issue of mental images. It seems very likely that subjects had an image of the VCR control panel and simply "read off" the various controls or displays. Related to this, the instructional material was more functionally organized; type links such as "type of button" were not a large part of the instructional graph. Since question probes were a direct function of the instructional graph, type links were infrequently probed. Finally, in developing the conceptual graphs, we were very conservative in graphing only relationships specified in the question probe answers and not _implicit_ links. Many concepts were part of several nodes, however, it was our policy not to connect these nodes simply on the basis of a shared concept. Were this policy to be changed, it might prove difficult to determine how much an experimenter could legitimately add to the conceptual graphs.

It might be possible to combine other methods of inferring conceptual associations with the question probe technique. For example, free recall and reaction time are good methods for measuring associations between concepts. If their respective disadvantages could be overcome, these methods might be used in conjunction with question probes to enhance the completeness of the graphs. This enhancement would be important should the graphs be used to model sequential operations acting upon a graph during problem solving.

Implications for Training and Education

The work reported here is applicable to training and education in a number of respects. These applications fall under two categories, the research method and the research results. First, the question probe method can be used to externalize expert knowledge structures for instructional design (Moore &

Gordon, 1988). Gordon and Gill (1989) have successfully applied the question probe method to the knowledge elicitation process in expert system development. Using question probes with experts is especially effective if question probes and conceptual structure development are used iteratively.

The resulting graph can then be used for a variety of purposes. They make knowledge content and structure explicit, and therefore support and enhance an instructor's ability to design course content and organization. The structures can also be used to design course texts and computer-aided instruction modules. A particularly interesting possibility is the newly evolving use of hypertext in computer tutorial systems (e.g., Jaffe & Lynch, 1989). Given that hypertext is essentially a computerized associative network, conceptual structures could provide both organization and content for such systems.

Similarly, conceptual graphs can be used to evaluate existing documents and texts. The document is first translated into a conceptual graph which may be segmented depending on the size of the graph. An expert can review the graph for missing transitional links, missing explanatory or procedural nodes, and inconsistencies. Question probes elicit information from the subject matter expert to complete and clarify the graph. We are currently using these methods to develop an engineering text.

The question probe and graphing methods could also be a useful tool for educational research. As an example, researchers have noted that students sometimes retain prior misconceptions even when given accurate information in the classroom (Gauld, 1988; Siegler, 1983). Question probes could be used as a means of evaluating the impact of various instructional strategies and content on knowledge structures (Gill, Gordon, Moore, & Barbera, 1988).

Question probes can also be used as a diagnostic tool for students. In discussing the prevalent problem of misconceptions in the classroom, Siegler (1983) stated, "Once teachers possessed this information [knowledge of common misunderstandings], they could design tests to reveal not only which children did not understand a concept, but also the nature of each child's misunderstanding" (p. 638). Question probes can be used to identify the exact nature of misconceptions as well as critical missing information. The technique could be used either by the instructor or within computerized systems to augment problem solving diagnostic routines.

Finally, our experimental results speak to the general process of education as well. These analyses suggest that for topics typical of an academic environment, what students learn and how that knowledge is organized are factors critical to successful problem solving performance. While inferencing is accomplished by combining existing pieces of knowledge, there was little evidence for substantial abstract inferencing and reasoning. Consistent with other ongoing research (e.g., Schneider et al., 1989), our studies indicate that performance is a function of the amount of directly applicable domain knowledge. We cannot assume that if we teach students basic concepts, theories, and principles, the knowledge will be applied

44

later on in real-life situations.

Directions for Future Research

The work reported here sets the stage for research in several directions. We plan to evaluate a number of methods for improving the accuracy of the conceptual graphs. While the content and organization of the graphs were strongly validated by the recall data, there were indications of more "single concept" associations than were shown in the structures. This was partly due to our bias towards being conservative in developing the graphs, mapping only what was stated in question probes. One future goal is therefore to identify methods for adding these simple associations into the graphs and evaluate their role in problem solving.

By enhancing the degree to which the graphs reflect underlying associations, we will move toward our primary goal which is to develop and test a more specific model of knowledge structure search during problem solving. Such a model should predict not only what procedures are performed but the exact sequential nature of the procedures as well as the various ideas, hypotheses, instances retrieved, etc. that occur during the problem solving process.

Once the conceptual structures have been sufficiently well-defined to allow testing of search models, the paradigm can be broadened to study other related questions. For example, we found that subjects accessed the majority of their knowledge structures during problem solving. However, this could easily be due to the experimental conditions of the studies. Potts, St. John, & Kirson (1989) have recently shown that information presented under "instructional learning" conditions such as those used here tend to result in subjects' compartmentalizing the knowledge rather than integrating it into their previously existing knowledge. People performing problem solving tasks in a real world environment will often have larger relevant conceptual structures, and those structures will be spread throughout a network among structures irrelevant to the problem at hand. It is clear that a "complete access" assumption will not be applicable under those circumstances. Thus, it will be necessary to study problem solving under more real-world circumstances to determine the process and limitations of the search mechanism. The methodologies developed in this research will allow a wide range of other questions to be addressed as well, including the effects of stress, secondary tasks, and other deleterious factors influencing problem solving activity.

Acknowledgements

References

Adelson, B. (1984). When novices surpass experts: The difficulty of a task may increase with expertise. _Journal of Experimental Psychology: Learning, Memory, and Cognition_, _10_, 483-495.

Bousfield, W.A. (1953). The occurrence of clustering in the recall of randomly arranged associates. _Journal of General Psychology_, _49_, 229-240.

Chase, W.G., & Simon, H.A. (1973). Perception in chess. _Cognitive Psychology_, _4_, 55-81.

Chi, M.T.H. (1978). Knowledge structures and memory development. In R.S. Siegler (Ed.), _Children's thinking: What develops?_ (pp. 73-96). Hillsdale, NJ: Erlbaum.

Chi, M.T.H., Feltovich, P.J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. _Cognitive Science_, _5_, 121-152.

Chi, M.T.H., & Glaser, R. (1985). Problem-solving ability. In R.S. Sternberg (Ed.), _Human abilities: An information-processing approach_. New York: W.H. Freeman and Company.

Chiesi, H.L., Spilich, G.J., & Voss, J.F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. _Journal of Verbal Learning and Verbal Behavior_, _18_, 257-274.

deJong, T., & Ferguson-Hessler, M.G. (1986). Cognitive structures of good and poor novice problem solvers in physics. _Journal of Educational Psychology_, _78_, 279-288.

diSessa, A. (1983). Phenomenology and the evolution of intuition. In D. Gentner and A.L. Stevens (Eds.), _Mental Models_. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Ericsson, K.A. & Simon, H.A. (1984). _Protocol analysis: Verbal reports as data_. Cambridge, Mass.: MIT press.

Evans, J.St.B.T. (1982). _The psychology of deductive reasoning_. London: Routledge & Kegan Paul.

Forbus, K. (1981). Qualitative reasoning about physical processes. _Proceedings of the Eighth IJCAI_, Vancouver, B.C., pp. 326-30.

Gauld, C.F. (1988). The cognitive context of pupils' alternative frameworks. _International Journal of Science Education_, _10_, 267-274.

Gill, R., Gordon, S., Moore, J., & Barbera, C. (1988). The use of conceptual structures in problem solving. <u>Proceedings, Annual Meeting of the American Society for Engineering Education</u>, <u>2</u>, 583-590.

Gordon, S., & Gill, R. (1989). Question probes: A structured method for eliciting declarative knowledge. <u>AI Applications in Natural Resource Management</u>, <u>3</u>, 13-20.

Graesser, A.C. & Clark, L.F. (1985). <u>Structures and procedures of implicit knowlege</u>. Norwood, N.J.: Ablex.

Graesser, A.C. & Goodman, S.M. (1985). Implicit knowledge, question answering, and the representation of expository text. In B.K. Britton & J.B. Black (Eds.), <u>Understanding expository text</u>. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Hinsley, D.A., Hayes, J.R., & Simon, H.A. (1978). From words to equations: Meaning and representation in algegra word problems. In P.A. Carpenter and M.A. Just (Eds.), <u>Cognitive processes in comprehension</u>. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Jaffe, C.C., & Lynch, P.J. (1989). Hypermedia for education in the life sciences. <u>Academic Computing</u>, September, 10-57.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. <u>Psychological Review</u>, <u>95</u>, 163-182.

Kintsch, W., & Greeno, J.G. (1985). Understanding and solving word arithmetic problems. <u>Psychological Review, 92</u>, 109-129.

Kolodner, J.L. & Simpson, R.L. (1986). Problem solving and dynamic memory. In J.L. Kolodner & C.K. Riesbeck (Eds.), <u>Experience, memory, and reasoning</u>. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Larkin, J.H. (1983). The role of problem representation in physics. In D. Gentner & A.L. Stevens (Eds.), <u>Mental models</u>. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lesgold, A. (1988). Problem solving. In R.J. Sternberg and E.E. Smith (Eds.), <u>The psychology of human thought</u>. Cambridge: Cambridge University Press.

Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing X-ray pictures. In M. Chi, R. Glaser and M. Farr (Eds.), <u>The nature of expertise</u>. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Moore, J., & Gordon, S.E. (1988). Conceptual graphs as instructional tools. _Proceedings, 32nd Annual Meeting of the Human Factors Society_, _2_, 1289-1293.

Novick, L.R. (1988). Analogical transfer, problem similarity, and expertise. _Journal of Experimental Psychology: Learning, Memory, and Cognition_, _14_, 510-520.

Payne, J.W. (1980). Information processing theory: Some concepts and methods applied to decision research. In T.S. Wallsten (Ed.), _Cognitive processes in choice and decision behavior_. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Pirolli, P.L., & Anderson, J.R. (1984, November). _The rold of mental models in learning to program_. Paper presented at the twenty-fifth meeting of the Psychonomic Society, San Antonio, Texas.

Potts, G.R., St. John, M.F., & Kirson, D. (1989). Incorporating new information into existing world knowledge. _Cognitive Psychology_, _21_, 303-333.

Reed, S.K. (1987). A structure-mapping model for word problems. _Journal of Experimental Psychology: Learning, Memory, and Cognition_, _13_, 124-139.

Riley, M.S., Greeno, J.G., & Heller, J.I. (1981). Development of children's problem-solving ability in arithmetic. In J.P. Ginsburg (Ed.), _The development of Mathematical thinking_. New York: Academic Press.

Rumelhart, D. & Norman, D. (1981). Analogical processes in learning. In J. Anderson (Ed.), _Cognitive skills and their acquisition_. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schneider, W., Korkel, J, & Weinert, F.E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low-aptitude children. _Journal of Educational Psychology_, _81_, 3060312.

Schoenfeld, A.H., & Herrmann, D.J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. _Journal of Experimental Psychology: Learning, Memory, and Cognition_, _8_, 484-494.

Sembugamoorthy, V. & Chandrasekaran, B. (1986). Functional representation of devices and compilation of diagnostic problem-solving systems. In J.L. Kolodner & C.K. Riesbeck (Eds.), _Experience, memory, and reasoning_. Hillsdale, N.J.: Lawrence Erlbaum Associate.

Siegler, R. (1983). How knowledge influences learning. _American Scientist_, _71_, 631-638.

Silver, E.A. (1981). Recall of mathematical problem information: Solving related problems. _Journal for Research in Mathematics Education_, _12_, 54-64.

Sowa, J.F. (1984). _Conceptual structures: Information processing in mind and machine_. Reading, Mass: Addison-Wesley Publishing Company.

Spilich, G.J., Vesonder, G.T., Chiesi, H.L., & Voss, J.F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. _Journal of Verbal Learning and Verbal Behavior_, _18_, 275-290.

Stevens, A., Collins, A., & Goldin, S. (1979). Misconceptions in students' understanding. _International Journal of Man-Machine Studies_, _11_, 145-156.

van Dijk, T.A. & Kintsch, W.A. (1983). _Strategies of discourse comprehension_. New York: Academic Press.

Weisberg, R., DiCamillo, M, & Phillips, D. (1978). Transferring old associations to new situations: A nonautomatic process. _Journal of Verbal Learning and Verbal Behavior_, _17_, 219-228.

Williams, M.D., Hollan, J., & Stevens, A. (1983). Human reasoning about a simple physical system. In D. Gentner and A. Stevens (Eds.), _Mental Models_, Hillsdale, N.J.: Lawrence Erlbaum Associates.

Engineering Mechanics Graph and Question Probes

     The instructional videotape consisted of an explanation of "how to mathematically model vectors". This appendix lists all of the main node-link-node segments contained in the graph of the instructional tape. These segments are listed in the left column. A node is listed followed by associated link-node pairs (e.g., `vector-IS-mathematical concept' is one node-link-node segment in the graph). To the right of the segments are the question probes that were developed to elicit information related to the concepts. Question probes in () denote question probes that were used in Experiment 1 but removed for Experiment 2. Questions or parts of questions in boldface type were additions for Experiment 2. Items in the graph segments denoted by [] represent verbal descriptions of material that were actually visual graphics in the instructional tape.

Item in Instruction	Question probe
Topic of Tape	
IS vectors	What was the topic of the
IS write vector(s)	video-tape?
Vector(s)	
IS mathematical concept	What is a vector?
IS mathematical quantity	
IS new topic	
IS arrow	
SYMBOL letter with line	What is the **algebraic or**
over it	**equation** symbol for a vector?
EQUATION $\overline{V}$ = Mag $\overline{U}v$	What is the equation for
Magnitude	a vector?
times	
$\overline{U}v$	
REASON solve problems	Why do we use vectors?
PROPERTY direction	What are the properties of
PROPERTY magnitude	vectors?
PROPERTY mathematically	
represented	
PROPERTY graphical representation	
PROPERTY location	
PROPERTY write vectors	
PROPERTY has a unit vector	
in same direction	
SUBTYPE unit vector	(What are the types of
SUBTYPE basic vectors	vectors?)
Direction	
IS direction vector is pointing	What is vector direction?
IS way arrowhead points	

INSTANCE due south   What is an example of vector
INSTANCE (example problems)   direction?

Magnitude
    IS length of line   What is vector magnitude?
    IS how big vector is
    IS conventional *number*
    IS size of vector
    INSTANCE 55 mph   What is an example of vector
    INSTANCE example problems   magnitude?
    SUBSET Magnitude in 2-D
    SUBSET Magnitude in 3-D
    SUBSET Magnitude of unit vector
    SYMBOL line on each side   What is the **algebraic or**
    SYMBOL $|\bar{V}|$ **equation** symbol for magnitude?

Magnitude in 2-D
    EQUATION $|\bar{V}| = x^2 + y^2$ What is the equation for
                                            vector magnitude in 2-D?

Magnitude in 3-D
    EQUATION $|\bar{V}| = x^2 + y^2 + z^2$ What is the equation for
                                            vector magnitude in 3-D?

Graphical representation
    IS arrow   **What does a graphically drawn**
    MANNER draw a directed line   **vector look like?**
           segment
    PROPERTY drawn on coordinate   (How do you represent
        axis   vectors graphically?)
    INSTANCE [picture of graph]

Drawn on coordinate axis
    INSTANCE [picture of graph]

Coordinate axis
    SUBTYPE 2-D coordinate axis   What were the two different
    SUBTYPE 3-D coordinate axis   kinds of coordinate axes
    INSTANCE [picture with 3 arrows]   **used to draw vectors?**

2-D Coordinate Axis
    IS REPRESENTED BY [picture   (How is a 2-D coordinate axis
                 with two arrows]   represented?)

    PART X axis   What are the parts of a 2-D
    PART Y axis   coordinate axis?
    PART origin

3-D Coordinate Axis
    IS REPRESENTED BY [picture with   (How is a 3-D coordinate axis
                 three arrows]   represented?)

    PART X axis   What are the parts of a 3-D
    PART Y axis   coordinate axis?
    PART Z axis
    PART origin

51

```
Origin
    IS where axes meet                      What is the origin?

X axis
    PROP line goes left and right           What is the X axis?
    PART Positive X axis
    PART Negative X axis

Y axis
    PROP line goes up and down              What is the Y axis?
    PART Positive Y axis
    PART Negative Y axis

Z axis
    PROP line goes in & out of paper        What is the Z axis?
    PART Positive Z axis
    PART Negative Z axis

Positive X axis
    PROP  line goes to right from           Where is the positive part
          origin                            of the X axis located?
                                            (How is the positive X axis
                                             represented?)

Negative X axis
    PROP line goes to left                  Where is the negative part
         from origin                        of the X axis located?
                                            (How is the negative X axis
                                             represented?)

Positive Y axis
    PROP line goes up from                  Where is the positive part
         origin                             of the Y axis located?
                                            (How is the positive Y axis
                                             represented?)

Negative Y axis
    PROP line goes down from                Where is the negative part
         origin                             of the Y axis located?
                                            (How is the negative Y axis
                                             represented?)

Positive Z axis
    PROP line comes out from                Where is the positive part
         origin                             of the Z axis located?
                                            (How is the positive Z axis
                                             represented?)

Negative Z axis
    PROP line goes into                     Where is the negative part
         page/screen from origin            of the Z axis located?
                                            (How is the negative Z axis
                                             represented?)


Basic Vectors
    TYPE i                                  What are the three basic
    TYPE j                                  vectors?
    TYPE k
    ENABLE write other vectors
```

Unit Vector
    IS any vector with magnitude
      of one
    PROPERTY magnitude
    PROPERTY length
    SYMBOL $\bar{U}v$

    PROPERTY has unit vector equation
    INST $\bar{i}$
    INST $\bar{j}$
    INST $\bar{k}$
    INST $-\bar{i}$
    INST $-\bar{j}$
    INST $-\bar{k}$

What is a unit vector?

What is the **algebraic or equation** symbol for unit vector? (What are the properties of a unit vector?)

What are some instances of a unit vector?

Magnitude of unit vector
    IS one
    SYMBOL $|\bar{U}v|$
    SUBTYPE Magnitude of unit
        vector in 2-D
    SUBTYPE Magnitude of unit
        vector in 3-D

What is the magnitude of a unit vector?

Magnitude of unit vector in 2-D
    EQUATION
$$1 = \sqrt{(\cos\theta x)^2 + (\cos\theta y)^2}$$

What is the equation for writing the magnitude of a unit vector in 2-D?

Magnitude of unit vector in 3-D
    EQUATION
$$1 = \sqrt{(\cos\theta x)^2 + (\cos\theta y)^2 + (\cos\theta z)^2}$$

What is the equation for writing the magnitude of a unit vector in 3-D?

$\bar{i}$
    IS a basic vector
    IS a unit vector
    PROPERTY points in direction
        of positive x axis
    PROPERTY magnitude is one
    PROPERTY arrow for vector can be
        "on" x line
    PROPERTY arrow for vector can be
        parallel to x line

What is $\bar{i}$?

What are the properties **or characteristics** of $\bar{i}$?

$\bar{j}$
    IS a basic vector
    IS a unit vector
    PROPERTY points in direction
        of positive y axis
    PROPERTY magnitude is one
    PROPERTY arrow for vector can be
        "on" y line
    PROPERTY arrow for vector can be
        parallel to y line

What is $\bar{j}$?

What are the properties **or characteristics** of $\bar{j}$?

$\bar{k}$

    IS a basic vector   What is $\bar{k}$?

    IS a unit vector

    PROPERTY points in direction of   What are the properties **or**

             positive z axis   **characteristics** of $\bar{k}$?

    PROPERTY magnitude is one

    PROPERTY arrow for vector can be

             "on" z line

    PROPERTY arrow for vector can be

             parallel to z line

$-\bar{i}$

    PROPERTY points in direction   What are the properties of $-\bar{i}$?

             of negative x axis

    PROPERTY points to the left

$-\bar{j}$

    PROPERTY points in direction   What are the properties of $-\bar{j}$?

             of negative y axis

    PROPERTY points downward

$-\bar{k}$

    PROPERTY points in direction   What are the properties of $-\bar{k}$?

             of negative z axis

    PROPERTY points back into screen

θx

    IS angle of x   What is theta x (θx)?

    IS angle between positive

        x axis and vector

$\text{Cos}\theta x$

    IS how far you went in   (What is cosine θx?)

        X direction

θy

    IS angle of y   What is theta y (θy)?

    IS angle between positive y

        axis and vector

$\text{Cos}\theta y$

    IS how far you went in Y   (What is cosine θy?)

        direction

θz

    IS angle of z   What is theta z (θz)?

    IS angle between positive z

        axis and vector

$\text{Cos}\theta z$

    IS how far you went in Z   (What is cosine θz?)

        direction

```
        PROPERTY number will be plus          [material presented in
              or minus                         an example problem and
                                               thus inadvertently
Number will be plus or minus                   missed during question
   CONS decide if number is                    probe development]
         plus or minus
Decide if number is plus or minus
   MANNER positive if vector pointing
           forward

Write vector(s)
   MANNER use vector equation
   PROCEDURE                                  What is the procedure for
      Find magnitude                          writing a vector?
    NEXT
      Write a unit vector
    NEXT
      Multiply magnitude
          times the unit vector
   PROPERTY independent of location

Write a unit vector
   MANNER Use unit vector equation
   SUBTYPE write a unit vector in 2-D
   SUBTYPE write a unit vector in 3-D

Unit vector equation
   SUBTYPE unit vector equation             (What are the types of
           for 2-D                            unit vector equations?)
   SUBTYPE unit vector equation
           for 3-D
```

Unit vector equation for 2-D
    EQUATION $\bar{U}v = \cos\theta x \bar{i}$ & $\cos\theta y \bar{j}$ What is the equation for
                                                writing a unit vector

```
    PROPERTY three conditions              in 2-D?

Three conditions
   IS vector starts at origin              What are the three conditions
   IS vector points away from              for using the unit vector
      origin                               equation?
   IS vector angles are measured
      from positive axes

Write a unit vector in 2-D
   MANNER use unit vector
           equation for 2-D
   PROCEDURE                               What is the procedure for
     Check three conditions                writing a unit vector in 2-D?
    NEXT
     If given angle θy
     Find θx for 2-D vector
    OR If given angle θx
     Find θy for 2-D vector
```

```
    NEXT
     Find cosine of angle θx
    AND
     Find cosine of angle θy
    NEXT
     Multiply cosθx times i
    AND
     Multiply cosθy times j
    NEXT
     Add together

Find θy for 2-D vector
   PROCEDURE                          What is the procedure for
      subtract θx from 90             finding θy in 2-D if given
                                      θx?
Find θx for 2-D vector
   PROCEDURE                          (What is the procedure for
      subtract θy from 90             finding θx in 2-D if given
                                      θy?)
Find cosine of angle θx
   PROCEDURE
      Calculate cosine of angle       What is the procedure for
                                      finding the cosine of an
Find cosine of angle θy               angle?
   PROCEDURE
      Calculate cosine of angle

Calculate cosine of angle
   PROCECURE
       Punch number on calculator
       Press cosine key

Unit vector equation for 3-D
   EQUATION                        _  What is the equation for
      Uv = cosθxī + cosθyj̄ + cosθzk  writing a unit vector in 3-D?
   PROPERTY three conditions

Write a unit vector in 3-D
   MANNER use unit vector
         equation for 3-D
   PROCEDURE                          What is the procedure for
      Check three conditions          writing a unit vector in 3-D?
   NEXT
      Find missing cos0
   NEXT              _
      Multiply cosθx times i
   AND              _
      Multiply cosθy times j
   AND              _
      Multiply cosθz times k
   NEXT
      Add together
```

```
Find missing cosθ (for 3-D)
      PROCEDURE                          What is the procedure for
        Get equation for magnitude       finding the missing cosθ
          of a unit vector in 3-D        in 3-D?
      NEXT
        Enter two known angles into
          equation
      NEXT
        Solve for missing cosθ

Vector location
    INST vector starts at origin
    INST vector doesn't start at origin
    INST vector points away from origin
    INST vector doesn't point away
        from origin
    PROPERTY angle is important
    PROPERTY location unimportant

Location unimportant
    REASON no reference to location      Why isn't location of a vector
    CONSEQUENCE can redraw vector        important?
            anywhere

Vector doesn't start at origin
    IS NOT vector starts at origin
    CONSEQUENCE redraw vector            What do you do if the vector
            starting from origin         doesn't start at the origin?

Redraw vector starting at origin
    MANNER move vector to origin

Move vector to origin
    PROP must keep angles the same

Vec. doesn't point away from origin
    IS NOT vector points away from
            origin
    CONSEQUENCE redraw vector            What do you do if the vector
            pointing away from           doesn't point away from the
            origin                       origin?

Redraw vector pointing away
from origin
    PROCEDURE
        slide through origin
      NEXT
        refigure the angles              What do you do after you slide
                                         the vector through the origin?

Refigure the angles
    PROCEDURE
        Look at opposite                 How do you refigure the angles
            angles to determine          after you slide the vector
            missing angles               through the origin?
```

57

NEXT
      Measure from positive axes

Vector measured from negative axis
    IS NOT measured from positive
          axis
      CONSEQUENCE re-measure angle   What is the consequence of
            from positive axis   the vector angle measurement
                                            being from the negative axis?

Re-measure angle from pos axis
    PROCEDURE   What is the procedure for
        measure clockwise from   re-measuring the angle from
            the positive axis   the positive axis instead of
        OR measure counterclockwise   the negative axis?
            from the positive axis
    PROPERTY doesn't matter if
            measure clockwise or
            counterclockwise

Doesn't matter if measure clockwise
or counterclockwise
    REASON cosine of angle measured   Why doesn't it matter if
            will be the same   you re-measure from the
                                            positive axis in either a
                                            clockwise or counterclockwise
                                            directions?

APPENDIX II

VCR Graph and Question Probes

        The instructional videotape consisted of an explanation of
how to use a JVC Hi-Fi VCR with flying head.   This appendix lists
a representative subset of the entire conceptual graph (which had
393 nodes in Experiment 3).   The segments are listed in the left
column as node-link-node pairs.   To the right of the segments are
associated question probes.   The question probes for this
particular segment of the graph were used in both Experiment 3
and 4.

Item in Instruction   Question Probe

Source select button
    PROP slides   What are the properties
    PROP located at lower left   or characteristics of
        corner of display panel   the source select button?
    PROP "source select"
    PROP positions

Positions
    SET tuner position   What are the source
    SET sc position   select positions?
    SET aux position

Tuner position
    PROP labeled "tuner"
    PROP on left side   Where is the tuner
                                                 position located?

    OUTCOME VCR records a TV channel   What happens when the
                                                 source select button
"tuner"   is in the TUNER position?
    PROP underneath source
        select button
    PROP on left side

Record a TV channel
    MANNER VCR selects a channel   How does the VCR record
    MANNER record normal audio track   a TV channel (that is,
        from TV   where does it record the
    MANNER record hi-fi audio track   different signals)?
        from TV

SC position
    PROP "sc"
    OUTCOME record normal   What happens when the
        audio channel from TV   source select button
    OUTCOME record hi-fi audio channel   is in the SC position?
        from aux input
    OUTCOME records TV video signal

"sc"
    IS simulcast   What does SC stand for?
    IS simultaneous broadcast
    PROP in middle   Where is the "sc"
    PROP underneath source   position located?
        select button
    INST record a concert

Aux position
    PROP "aux"
    PROP on right side
    OUTCOME records aux video signal   What happens when the
    OUTCOME records aux audio signal   source select button is
                                       in the AUX position?
"aux"
    PROP on right side   Where is the AUX
    PROP underneath source   position located?
        select button

Set record level(s)
    PROP can't set normal channel level   Which "record" audio
    PROP can set hi-fi record level   levels can be set?
    PROP important

Can't set normal channel level
    REASON normal channel always   Why can't you set the
        automatically controlled   normal audio channel?

Normal channel always autom. controlled
    MANNER VCR automatically adjusts

Important (t⌣ ⁔t record level)
    REASON don't want sound too low   Why is it important to
    REASON don't want sound too loud   set the "record" audio
                                       level?
Sound too low
    OUTC lots of noise and hiss   What happens if the
    OUTC difficult to hear   sound is set too low?

Sound too loud
    OUTC sound will distort   What happens if the
                                     sound is set too high?

Set hi-fi record level
    MANNER automatic mode   What are the two ways
    MANNER manual mode   to set the hi-fi
    MANNER ALC switch   "record" level?

Automatic mode
    PROP VCR automatically adjusts   What is the automatic
                                     mode for the record
                                     level?

Manual mode
    PROP you control level
    MANNER ALC switch
    MANNER sound control switches

ALC switch
    PROP above and left
        of source switch
    PROP positions
    PROP can be on
    PROP can be off

Positions
    SET to left
    SET to right
To left
    OUTCOME ALC is off

To right
    OUTCOME ALC is on

ALC is on
    OUTCOME VCR automatically
          adjusts level

ALC is off
    OUTCOME you control level
    OUTCOME can tell sound level

You control level
    MANNER sound control switches

Sound control switches
    PROP Two slide switches
    SET left channel switch
    SET right channel switch
    PROP settings
    PROP located in center top
        of panel
    PROP move to left
    PROP move to right

Left channel switch
    PROP on top

Right channel switch
    PROP on bottom

Settings
    INST on left end
    INST on right end

How do you set the
sound level using the
manual mode?

What are the properties
or characteristics of
the ALC switch?

What are the positions
of the ALC switch?

What happens if the ALC
switch is on?

What happens if the ALC
switch is off?

How do you control the
sound level?

What are the properties
or characteristics of
the sound control switches?

Where is the left channel
sound control switch?

Where is the right channel
sound control switch?

What are the positions
or settings for the
sound control switches?

APPENDIX III

FREE RECALL CLUSTER SCORES

Each of the columns represents the sequence of "cluster" scores for a subject. Each number represents the number of conceptual graph links separating a free recall statement from the previous one. Zeros signify that two propositions given in free recall were clustered or directly contiguous on the question probe graph. An "S" (for "skip") indicates that the statement went to a new section of the conceptual graph, not connected by any means to the previous one. The Asterisks (*) indicate that the item contained either a new node or link not found on the conceptual graph. The numerical score was based on the part of the statement that was in the graph.

Free Recall Sequence	Subject											
	1	2	3	4	5	6	7	8	9	10	11	12
1												
2	0	3	0	0*	S	0	0*	0	0	0	0	0
3	0	0	1*	0*	0	1	0	0	0	0	0	0
4	0	1	0	1	0	0*	0*	8	0	1	S	1*
5	0	2	0	0	S	0	0	0	1	0	0	1
6	0*	2	0	0	S*	1	5	0	0	7	S	0
7	2	0	2*	0	0*	0	0	0	S	1*	0	0
8	0	7	0	0*	0	1	1	2	3	0	0	1
9	0	1*	0	0	S	0	0	0*	S	4	1	S
10	0	0*	0	0	0	1	0	0	S	0	0	S
11	0	4	0*	S*	0	1	0	0	3	0	S	0
12	0	0	3	0	1	0	1	0	1	0	0	2
13	1	3	0	1	0	5	0	0	S	2	0	4
14	0*	4	0	0	1	0	S	0	2	0	0	0
15	0	0	0	S	0	0	0	S	0	S*	7*	S
16	0	S	S	0	S	2	S	0*	S	0	0	2
17	0	1	0*	0	0	0	0	S	0	0	0	S
18	2*	0	1	S*	0*	2*	0	0	1	S	0	0
19	1	5	0	0*	0*	0*	1	0	0	0	0	0
20	0	0	0	1	0	S*	0	0	0	0	0	3
21	S	1	S	0	S	0	4	0	S	5	S	0
22	0*	0	0	0	0*	2	S*	0	1	2	7	S
23	0	S	0	2	S	S	0*	0	0	0	2	4
24	0	2	0	0	0	0		0	3	0	S	2
25	S	0*	S	1*	0	0		0*	0	0	S	0
26	S	2	0	S	1*	2		0*	S	S	0	1*
27	0	0	1*	S	2	0		S*	0	0	0	S
28	0	2	0	8	0	3*		0*	0	0	2	0
29	2*	0*	0*	S	0*	S*		0	1	2	0	
30	2	1*	0*	0	0*	0		0	0	2	4	
31	0*	0	0	S	2	0*		0	1	0	0	
32	0	S	0*	0	1	4		0	0	2	4	
33	0	S	S	0	0	0		S	S	3	0	
34	1	5	2	0	0	0		0	0	0	0	

					Subject							
	1	2	3	4	5	6	7	8	9	10	11	12
35	S	0		0	2*	2		S	0	2	0	
36	0*			0	0	0*		0		0		
37	S			0	0	0		S*		5		
38	0			0	0	0		0*		0		
39	2			S	0	1		0*		0		
40	2			0	0	0		S		0		
41				0	0	0		0		1		
42					0	S		S		1		
43					1	0*				0		
44					0	0				S		
45					1	1				0		
46					0	S				2		
47					0	0*				S		
48					0	0				6		
49					S*	S				S*		
50					0	1				0*		
51					0	0				1		
52					3					0		
53					5					1		
54					0					2		
55					0					1		
56					S*					0		
57					0*					0		
58					1					0		
59					0					7		
60					0					5		
61					S					0		
62					0					S		
63					1					0		
64					0					1		
65					4					0		
66					4					3		
67					S*							
68					0							
69					S							